Chapter 4

The Elimination Procedures for the Competition Number and the Phylogeny Number

4.1 Introduction

Given an acyclic digraph D, the competition graph $C(D)$ is defined to be the undirected graph with $V(D)$ as its vertex set and where vertices x and y are adjacent if there exists another vertex z such that the arcs (x, z) and (y, z) are both present in D. Competition graphs were introduced by Cohen [1] to study ecosystems. The vertices of an acyclic digraph D, known as a food web, represent species, and the arc (x, z) indicates that z is a prey of x. An edge exists between two vertices x and y in $C(D)$ if and only if x and y have a common prey. In addition to ecology, competition graphs have also found application in studying communication over noisy channels, interfering radio transmissions, and models of complex economic and energy problems – see the discussions in Raychaudhuri and Roberts [9] and Roberts [12]. Lundgren [6], Roberts [10], and Kim [3] survey the extensive literature of competition graphs.

In [11] Roberts noted that for any graph G, G along with r isolated vertices is the competition graph of some acyclic digraph if r is sufficiently large. The competition number $k(G)$ is defined to be the least such r. In general, determining the competition number of a graph is difficult: Opsut [7] showed that this problem is NP-complete. Kim and Roberts in [11] and [4] have determined the competition number of graphs with 0, 1, and 2 triangles, but for few other graph classes is the competition number known. As another approach, Roberts considered using an elimination procedure to calculate $k(G)$. An elimination procedure takes as input G and an ordering $\mathcal{O} = v_1, \ldots, v_n$ of the vertices of G and produces an acyclic digraph D such that $C(D) = G \cup I_r$; that is, the competition graph of D is G

along with r isolated vertices. The procedure "eliminates" each vertex in order by ensuring that all of the edges incident on the vertex will appear in $C(D)$. The goal is to create an elimination procedure that for some ordering $\mathcal{O}$ outputs an acyclic digraph D where $|V(D) \setminus V(G)| = k(G)$.

Elimination procedures which seek to determine a graph-theoretical parameter through step-wise elimination of vertices have various applications in graph theory. A common example are algorithms for determining if a graph is chordal by finding perfect elimination orders such that each vertex is "simplicial" in the graph of remaining vertices. Roberts [11] was led to consider an elimination procedure for the competition number through variants of perfect elimination used by Parter [8], Rose [15], and Golumbic [2] in connection with numerical analysis. Here elimination procedures are used to find a good order for eliminating variables during Gaussian elimination of a matrix.

Opsut [7] found an example of a graph G where Roberts' original elimination procedure does not calculate the competition number $k(G)$, thus giving a counterexample to Roberts' conjecture that the procedure always calculates $k(G)$. Kim and Roberts [5] then modified the elimination procedure and asked whether their modified procedure works for all graphs. They were able to show that the modified version calculates the competition number for a large class of graphs, the so-called "kite-free" graphs.

In this chapter, we present a new, simpler proof of Kim and Roberts' theorem that their elimination procedure calculates the competition number for kite-free graphs. We also present a graph L where Kim and Roberts' elimination procedure does not always calculate the competition number, in the following sense: for each order $\mathcal{O}$ of vertices of L, the elimination procedure can produce an acyclic digraph with more than $k(L)$ additional vertices.

Phylogeny graphs are related to competition graphs, and were introduced by Roberts and Sheng [13] from an idealized model for reconstructing phylogenetic trees. Given an acyclic digraph D, the phylogeny graph $P(D)$ is defined to be the undirected graph with $V(D)$ as its vertex set and with adjacencies as follows: two vertices x and y are adjacent if one of the arcs (x, y) or (y, x) is present in D, or if there exists another vertex z such that the arcs (x, z) and (y, z) are both present in D. Roberts and Sheng noted that for any simple graph G, G is an induced subgraph of $P(D)$ for some acyclic digraph D. The phylogeny number $p(G)$ is the least number r such that D has $|V(G)| + r$ vertices. Determining the phylogeny number of a graph was shown by Roberts and Sheng [13] to be NP-complete, and they also determined the phylogeny number for connected graphs with 0, 1, and 2 triangles [14].

Phylogenetic tree reconstruction deals with establishing evolutionary relationships between different species. A phylogenetic tree is a rooted directed tree, where the species are vertices and an arc (x, z) indicates that z is a direct ancestor of x. Given a set of species and a measure of similarity between each pair of species, we wish to create a phylogenetic tree where similar species are closely related. The concept of "closely related" can be defined in many ways, and Roberts and Sheng choose the tree metric where the distance between x and y is the shortest distance to a common ancestor. If the similarity measure is purely "similar" or "not similar," then we can encode the similarity relationship in a graph G, where two vertices are adjacent if and only if the corresponding species are similar. If we relax the condition of finding a tree, then finding a phylogenetic acyclic digraph D for the set of species turns out to be the same as finding an acyclic digraph D such that $P(D) = G$. The number of assumptions made in defining phylogeny graphs considerably removes the concept from the original biological motivation. However, phylogeny graphs give a starting point for studying phylogenetic tree reconstruction, and the related competition graphs give rise to many interesting mathematical techniques and questions for phylogeny graphs. One question that we address in section 4.5 is the construction and analysis of an elimination procedure. For a survey on competition graphs and phylogeny graphs and their relation, see Roberts [10].

Both of the results mentioned above for competition numbers carry over to phylogeny numbers: we present an analogous elimination procedure for the phylogeny number, show that it calculates the phylogeny number for kite-free graphs, and show that the procedure does not calculate the phylogeny number for the graph L. The phylogeny case is simpler than the competition case because of how vertices are "handled" when eliminated, and allows greater insight into the concepts underlying these results.

We mention here that there are many other variants of the competition number besides the phylogeny number. The common enemy graph of a digraph $D = (V, A)$ is the graph with vertex set V and where vertices x and y are adjacent if and only if there is a vertex a in D such that (a, x) and (a, y) are arcs of D (a is a common enemy or predator of x and y). In the niche graph of D, x and y are adjacent if and only if there is a vertex a such that (a, x) and (a, y) are arcs of D *or* there is a vertex b such that (x, b) and (y, b) are arcs of D (x and y *either* have a common predator a *or* a common prey b). In the competition-common enemy graph of D, x and y are adjacent if and only if there is a vertex a such that (a, x) and (a, y) are arcs of D *and* there is a vertex b such that (x, b) and (y, b) are arcs of D (x and y have *both* a common prey *and* a common predator). The common enemy number, niche number, and competition-common enemy number of a graph G can be defined analogously to competition number, and we refer the reader to [10] for a survey of results about these graph parameters. No one has considered elimination procedures for these parameters yet, and it would be interesting to see what results about the elimination procedures for the competition number and the phylogeny number carry over.

Note that the focus of creating elimination procedures is not on efficiency, since calculating the competition number or the phylogeny number with an elimination procedure requires $n!$ runs (one for each ordering of the vertices). As mentioned above, calculating both the competition number and the phylogeny number have been shown to be NP-complete. Instead, the focus is on whether an elimination procedure *could* be created that exactly calculates the relevant number. This is interesting both for historical reasons and because many of the practical examples are relatively small, exactness is sometimes more important than efficiency. Our results provide a better understanding of why Kim and Roberts' procedure is exact for kite-free graphs, and the counterexample suggests that creating an elimination procedure that is exact for all graphs might be much more difficult than originally thought.

In this chapter, the graph G that we wish to calculate the competition or phylogeny number of need not be connected. For convenience, we will sometimes also describe a subgraph H of a graph G only as "consisting of" certain edges of G. It is understood that H has no isolated vertices: the vertices of H are only the endpoints of edges in H.

4.2 The Elimination Procedure for the Competition Number

We will first formalize our definitions and describe Kim and Roberts' elimination procedure using our terminology; however, its workings are the same as the elimination procedure described in [5].

Definition 4.1. Let $D = (V, A)$ be an acyclic digraph. The *competition graph* $C(D)$ is a simple graph with vertex set V where two vertices x and y are adjacent in $C(D)$ if there exists a vertex z such that both (x, z) and (y, z) are arcs in D. From the ecological origins of competition graphs, z is known as a *prey* of x if (x, z) is an arc of D.

Definition 4.2. For a simple graph G, the *competition number* $k(G)$ is the least number r such that there exists an acyclic digraph D on $|V(G)|+r$ vertices where $C(D)$ is G along with r isolated vertices.

Before presenting the formal description of the elimination procedure for the competition number, we first give an informal description. Given a graph G and an ordering $\mathcal{O} = v_1, \ldots, v_n$ of the vertices of G, we eliminate each vertex iteratively, in the process building up an acyclic digraph D with the desired properties. When eliminating vertex v_i, we "cover" every edge incident to v_i that has not been covered in a previous iteration. By "covering" an edge e, we mean that the appropriate arcs and possibly vertices have been added to D so that e is an edge in $C(D)$. The subgraph G_i is a spanning subgraph of G that contains the edges of G that have not been covered in an iteration prior to the i^{th} iteration. The subgraph G'_i consists of the edges of G_i that are incident on v_i, and so the edges of G'_i must be covered in the i^{th} iteration. Cliques are used to maximize the coverage of G'_i using the least number of added vertices. If C is a clique, then by adding arcs in D_i from the vertices of C to a common vertex x, all of the edges in C appear in $C(D_i)$. Thus, all of the vertices in C are "preying" on the same "species" x, and hence competing with each other.

The improvement of Kim and Roberts' modified elimination procedure over Roberts' original procedure was in recognizing that the edges in G'_i are the *only* edges that must be covered in the i^{th} iteration. For choosing the cliques they utilize the subgraph H_i consisting of the edges from v_i to vertices of higher index. The cliques covering G'_i are chosen from H_i, even though some of the edges in H_i might already be covered. By using maximal cliques of H_i, possibly more uncovered edges that are not in G'_i will be covered.

Definition 4.3. Let $E_G(v)$ denote the subgraph of G with vertex set $N_G[v]$ and containing only those edges of G incident to the vertex v.

The Kim-Roberts Elimination Procedure for the Competition Number[1]

Input: A graph G and an ordering $\mathcal{O} = v_1, v_2, \ldots, v_n$ of the vertices of G.

Output: An acyclic digraph $D := D_n$ such that $C(D)$ is G with some additional isolated vertices.

Initialization: Set D_0 to the digraph with vertex set $V(G)$ and no arcs. D_i is an acyclic digraph constructed at the i^{th} iteration.

Set $G_1 := G$. G_i is a spanning subgraph of G that contains the edges of G that do not appear in $C(D_{i-1})$.

Set $S_1 := \emptyset$. S_i is a set of vertices available as prey.

$\boldsymbol{i^{\text{th}}}$ Iteration, $\boldsymbol{i = 1, \ldots, n}$: Set G'_i to $E_{G_i}(v_i)$, and set H_i to the subgraph of G induced by $\{v_i\} \cup \{v_j : j > i \text{ and } v_j \in N_G(v_i)\}$. Let $\mathcal{E}_i = \{C_1, \ldots, C_k\}$ be a minimum size edge covering of G'_i by maximal cliques of H_i, ordered arbitrarily. Form G_{i+1} from G_i by removing the edges of C_j from G_i for all j.

Form the digraph D_i by adding vertices and arcs to D_{i-1} as follows: Pick k distinct vertices $u_1, \ldots, u_k$ from S_i. If $|S_i| < k$, then add $k - |S_i|$ additional vertices $u_{k-|S_i|}, \ldots, u_k$ to D_i. For each clique $C_j \in \mathcal{E}_i$, add the arcs (w, u_j) to D_i for each $w \in C_j$. Form S_{i+1} by $S_{i+1} := (S_i \setminus \{u_1, \ldots, u_k\}) \cup \{v_i\}$.

Remark 4.4. Note that finding a minimum size edge covering of G'_i by maximal cliques of H_i is equivalent to finding a minimum size vertex covering by maximal cliques of $H_i \setminus \{v_i\}$

[1]If we refer simply to "the elimination procedure for the competition number," then we mean the Kim-Roberts elimination procedure.

of the subgraph induced by $N_{G_i}(v_i)$. The transformation between these procedures is as follows: For each clique C_j in $\mathcal{E}_i = \{C_1, \ldots, C_k\}$, set $\widehat{C}_j = C_j \setminus \{v_i\}$. Then $\widehat{\mathcal{E}}_i = \{\widehat{C}_1, \ldots, \widehat{C}_k\}$ is a minimum size vertex cover of $N_{G_i}(v_i)$ by maximal cliques of $H_i \setminus \{v_i\}$ if and only if $\mathcal{E}_i$ is a minimum size edge cover of G'_i by maximal cliques of H_i.

To help analyze the workings of the elimination procedure, we now introduce a more generalized elimination procedure. In the generalized elimination procedure, a clique cover of the entire graph G is given, and from this covering and the order of the vertices we construct D.

The Generalized Elimination Procedure for the Competition Number

Input: A graph G, an ordering $\mathcal{O} = v_1, v_2, \ldots, v_n$ of the vertices of G, and an edge clique covering $\mathcal{G}$ of G.

Output: An acyclic digraph $D := D_n$ such that $C(D)$ is G with some additional isolated vertices.

Initialization: Set D_0 to the digraph with vertices $V(G)$ and no arcs. D_i is an acyclic digraph constructed at the i^{th} iteration.

Set $S_1 := \emptyset$. S_i is a set of vertices available as prey.

i^{th} Iteration, $i = 1, \ldots, n$: Let $\mathcal{G}_i = \{C_1, \ldots, C_k\}$ be the subset of $\mathcal{G}$ where for each $C_j \in \mathcal{G}_i$, v_i is the vertex in C_j of least index. Order $\mathcal{G}_i$ arbitrarily.

Form the digraph D_i by adding vertices and arcs to D_{i-1} as follows: Pick k distinct vertices $u_1, \ldots, u_k$ from S_i. If $|S_i| < k$, then add $k - |S_i|$ additional vertices $u_{k-|S_i|}, \ldots, u_k$ to D_i. For each clique $C_j \in \mathcal{G}_i$, add the arcs (w, u_j) to D_i for each $w \in C_j$. Form S_{i+1} by $S_{i+1} := (S_i \setminus \{u_1, \ldots, u_k\}) \cup \{v_i\}$.

The following proposition is Proposition 1 from [5], if we note that though the proof is worded only for the elimination procedure, it also applies to the generalized elimination procedure.

Proposition 4.5. *The generalized elimination procedure for the competition number produces an acyclic digraph D where $C(D)$ is G along with some additional isolated vertices.*

We now show that the Kim-Roberts elimination procedure is a special case of the generalized elimination procedure.

Lemma 4.6. *Let $\mathcal{E}_i$ be the edge coverings generated by the Kim-Roberts elimination procedure for a graph G and a vertex ordering $\mathcal{O} = v_1, \ldots, v_n$. Then the set $\mathcal{E} = \bigcup_{i=1}^{n} \mathcal{E}_i$ is an edge clique covering of G.*

Proof. Since each $\mathcal{E}_i$ is chosen to be a set of cliques of H_i and H_i is a subgraph of G, $\mathcal{E}$ is a set of cliques of G. We now show that $\mathcal{E}$ covers all the edges of G. Let $\{v_k, v_\ell\}$ be an edge in G, where $k < \ell$. Suppose that $\{v_k, v_\ell\}$ is not an edge in any clique of $\bigcup_{i=1}^{k-1} \mathcal{E}_i$. Then G_k contains the edge $\{v_k, v_\ell\}$, as does G'_k. Since $\mathcal{E}_k$ is an edge clique covering of G'_k, there will exist a clique $C_j \in \mathcal{E}_k$ that contains $\{v_k, v_\ell\}$. Therefore, $\mathcal{E} = \bigcup_{i=1}^{n} \mathcal{E}_i$ is an edge clique covering of G. □

Proposition 4.7. *Let G be a graph and $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G. Then the number of vertices added to the digraph produced by the Kim-Roberts elimination procedure is the number of vertices added to the digraph produced by the generalized elimination procedure if the edge clique covering $\mathcal{G}$ is chosen to be $\mathcal{E}$ as defined in Lemma 4.6.*

Proof. The proposition follows from Lemma 4.6 and the observation that the subsets $\mathcal{G}_i$ used in the generalized elimination procedure are exactly the subsets $\mathcal{E}_i$ used in the elimination procedure. □

Furthermore, if in each i^{th} iteration the same clique C_1 is chosen from $\mathcal{G}_i = \mathcal{E}_i$, then the two digraphs are isomorphic.

In order to analyze the number of additional vertices needed by the elimination procedure to construct D, we would like a formula expressing this number in terms of the cliques chosen. We will give such a formula and show its correctness via the generalized elimination procedure.

Definition 4.8. The *elimination number* $M(G, \mathcal{O}, \mathcal{G})$ of a graph G, an ordering $\mathcal{O}$ of the vertices, and an edge clique covering $\mathcal{G}$ is the number of vertices added to G so that $C(D)$

is $G \cup I_{M(G,\mathcal{O},\mathcal{G})}$, where D is the digraph produced by the generalized elimination procedure for the competition number with G, $\mathcal{O}$, and $\mathcal{G}$ as inputs.

Definition 4.9. Let G be a graph, $\mathcal{O} = v_1, v_2, \ldots, v_n$ be an ordering of the vertices of G, and $\mathcal{G} = \{C_1, C_2, \ldots, C_k\}$ be an edge clique covering of G. For each vertex v_i, let $\mathcal{G}_i$ be the subset of $\mathcal{G}$ where for each $C_j \in \mathcal{G}_i$, v_i is the vertex in C_j of least index. Recursively define the integer sequences $\{a_i^{\mathcal{G}}\}_{i=0}^n$ and $\{b_i^{\mathcal{G}}\}_{i=1}^n$ by

$$\begin{aligned} a_0^{\mathcal{G}} &= 0, \\ b_i^{\mathcal{G}} &= \max\{|\mathcal{G}_i| - a_{i-1}^{\mathcal{G}}, 0\}, \\ a_i^{\mathcal{G}} &= a_{i-1}^{\mathcal{G}} - (|\mathcal{G}_i| - b_i^{\mathcal{G}}) + 1. \end{aligned}$$

Define

$$h_G(\mathcal{G}, \mathcal{O}) = \sum_{i=1}^{n} b_i^{\mathcal{G}}.$$

Note that $a_i^{\mathcal{G}}$ is the number $|S_i|$ of available prey at the end of the i^{th} iteration, and $b_i^{\mathcal{G}}$ is the number of new vertices added to D_i in the i^{th} iteration.

Lemma 4.10. *Let G be a graph, $\mathcal{O} = v_1, v_2, \ldots, v_n$ be an ordering of the vertices of G, and $\mathcal{G} = \{C_1, C_2, \ldots, C_k\}$ be an edge clique covering of G. Then $M(G, \mathcal{O}, \mathcal{G}) = h_G(\mathcal{G}, \mathcal{O})$.*

Proof. Note that $\mathcal{G}_i$ is defined in exactly the same way in both the generalized elimination procedure and in Definition 4.9. Summing $b_i^{\mathcal{G}}$ over all iterations, we get

$$M(G, \mathcal{O}, \mathcal{G}) = |V(D) \setminus V(G)| = \sum_{i=1}^{n} b_i^{\mathcal{G}} = h_G(\mathcal{G}, \mathcal{O}). \qquad \square$$

By taking a minimum over all edge clique covers and vertex orders of G, we can use $h_G(\mathcal{G}, \mathcal{O})$ to calculate the competition number of G.

Lemma 4.11. *For a graph G, the competition number $k(G)$ equals $\min_{\mathcal{G}} \min_{\mathcal{O}} h_G(\mathcal{G}, \mathcal{O})$, where $\mathcal{G}$ ranges over all edge clique coverings of G and $\mathcal{O}$ ranges over all orderings of the vertices of G.*

Proof. Let $\mathcal{G} = \{C_1, C_2, \ldots, C_k\}$ be an edge clique covering of G and $\mathcal{O} = v_1, v_2, \ldots, v_n$ be an ordering of the vertices of G. By Proposition 4.5, the generalized elimination procedure produces an acyclic digraph D such that $C(D)$ is G with some additional isolated vertices. By Lemma 4.10, $M(G, \mathcal{O}, \mathcal{G}) = h_G(\mathcal{G}, \mathcal{O})$, and so $k(G) \leq \min_{\mathcal{G}} \min_{\mathcal{O}} h_G(\mathcal{G}, \mathcal{O})$.

Now let F be an acyclic digraph that attains the competition number for G; that is, $C(F)$ is G along with isolated vertices and $|V(F) \setminus V(G)| = k(G)$. Let $\mathcal{O} = v_1, v_2, \ldots, v_n$ be an ordering of the vertices of G such that if (v_ℓ, v_k) is an arc in F, then $k < \ell$. Such an ordering exists because F is acyclic. We construct an edge clique covering $\mathcal{G}$ of G from F as follows: For a vertex $v_i \in V(G)$, $N_F^{\text{in}}(v_i)$ induces a clique in G, and for a vertex $b \in V(F) \setminus V(G)$, $N_F^{\text{in}}(b)$ induces a clique in G. Since having arcs into a common prey is the only way edges can be present in G, $\mathcal{G}$ is an edge clique cover of G. Now observe that the digraph D produced by the generalized elimination procedure with $\mathcal{G}$ and $\mathcal{O}$ has the same number of vertices as F. In fact, if the appropriate $u_1, \ldots, u_k$ are chosen from S_i, then D is isomorphic to F.

Therefore, $k(G) = |V(F) \setminus V(G)| = |V(D) \setminus V(G)| = M(G, \mathcal{O}, \mathcal{G}) \geq \min_{\mathcal{G}} \min_{\mathcal{O}} h_G(\mathcal{G}, \mathcal{O})$, and so $k(G) = \min_{\mathcal{G}} \min_{\mathcal{O}} h_G(\mathcal{G}, \mathcal{O})$. □

We now specialize the definition of the elimination number.

Definition 4.12. Given a graph G and an ordering $\mathcal{O}$, let $\mathcal{E} = \mathcal{E}(G, \mathcal{O}) = \{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n\}$ be edge clique coverings obtained during the Kim-Roberts elimination procedure. Of course, the notation is ambiguous since the way to choose the $\mathcal{E}_i$ is not completely specified in the procedure. The *elimination number* $M(G)$ is the minimum of $M(G, \mathcal{O}, \mathcal{E})$ over all orders $\mathcal{O}$ and some $\mathcal{E}$ obtained when using $\mathcal{O}$. Kim and Roberts show that for certain classes of graphs, if $M(G)$ is this minimum and is attained for $\mathcal{O}$ and some $\mathcal{E}$, then it is attained for the same $\mathcal{O}$ and any $\mathcal{E}$ corresponding to $\mathcal{O}$. If this is the case, $M(G)$ is unambiguously defined.

The determination of necessary and sufficient conditions for $M(G)$ to be unambiguously defined is an interesting open problem. Lemma 4.11 shows that there is always a "right" clique cover for each order such that the minimum over orders attains the competition number $k(G)$. Kim and Roberts show that there is a class of graphs, known as the kite-free graphs, for which $M(G)$ is unambiguously defined and equals $k(G)$. In the next section,

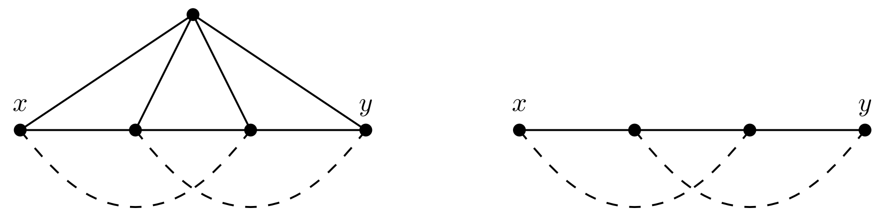

Figure 4.1: A kite and a kite-body.

we present a new and simpler proof of this result. Kim and Roberts also asked if $M(G)$ is unambiguously defined and equals $k(G)$ for all graphs. However, in section 4.4 we exhibit a graph L such that for each order $\mathcal{O}$ there is a choice of clique cover $\mathcal{E}_i$ in the Kim-Roberts elimination procedure such that $M(G, \mathcal{O}, \mathcal{E}) > k(G)$. This answers the Kim-Roberts question negatively.

4.3 Kite-free Graphs

In this section, we present a new and simpler proof of Kim and Roberts' theorem in [5] that their elimination procedure for competition numbers is exact for kite-free graphs.

Definition 4.13. A *kite* is the left configuration shown in Figure 4.1. In a kite, the solid edges must be present, and the dotted edges cannot be present. The edge between vertices x and y may or may not be present. A *kite-free* graph does not have a kite as a configuration, meaning that neither of the two graphs on five vertices that are kites are present as induced subgraphs. A *kite-body* is the right configuration shown in Figure 4.1. Again, the solid edges must be present, the dotted edges cannot be present, and the edge between x and y may or may not be present. Similarly, a *kite-body-free* graph does not have a kite-body as a configuration.

The following lemma is Lemma 3 from [5].

Lemma 4.14. *Let G be a kite-body-free graph, S a subset of $V(G)$, H an induced subgraph of G, and $C_1, C_2, \ldots, C_k$ a vertex cover of S using maximal cliques of H. If a subset T of S forms a clique in H, then T is contained in some C_ℓ.*

Lemma 4.15. *Let G be a kite-free graph and $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G. In the Kim-Roberts elimination procedure, an edge $\{v_j, v_k\}$ with $v_j, v_k \in N_G(v_i)$ appears in some clique of $\mathcal{E}_\ell$, where $\ell \leq i$.*

Proof. Suppose that $\{v_j, v_k\}$ with $v_j, v_k \in N_G(v_i)$ does not appear in any clique of $\mathcal{E}_\ell$, where $\ell < i$. Observe that in the elimination procedure all edges incident on a vertex v_ℓ are covered by $\bigcup_{r=1}^{\ell} \mathcal{E}_r$. Thus, $k > i$ and $j > i$, and so $\{v_j, v_k\}$ is an edge in $H_i \setminus \{v_i\}$. We now consider three different cases. Suppose that both v_j and v_k are in $N_{G_i}(v_i)$. Let $\mathcal{E}_i = \{C_1, \ldots, C_s\}$, and set $\widehat{C}_t = C_t \setminus \{v_i\}$. As stated in Remark 4.4, $\widehat{\mathcal{E}_i} = \{\widehat{C}_1, \ldots, \widehat{C}_s\}$ is a vertex cover of $N_{G_i}(v_i)$ by maximal cliques of $H_i \setminus \{v_i\}$. Since v_j and v_k are in $N_{G_i}(v_i)$, the edge $\{v_j, v_k\}$ forms a clique in $N_{G_i}(v_i)$. Since G is kite-free, $H_i \setminus \{v_i\}$ is kite-body-free. By Lemma 4.14, $\{v_j, v_k\}$ is a clique contained in some $\widehat{C}_\ell$, and so appears in clique C_ℓ of $\mathcal{E}_i$.

For the second case, suppose that neither v_j nor v_k is in $N_{G_i}(v_i)$. Since v_j is not in $N_{G_i}(v_i)$, the edge $\{v_j, v_i\}$ appears in some clique of $\mathcal{E}_\ell$, where $\ell < i$. Since v_k is also not in $N_{G_i}(v_i)$, then $\{v_k, v_i\}$ appears in some clique of $\mathcal{E}_m$, where $m < i$. If $\ell = m$, then by the first case applied to v_ℓ, $\{v_j, v_k\}$ appears in a clique of $\mathcal{E}_\ell = \mathcal{E}_m$. Thus, $\ell \neq m$. Since $\{v_j, v_k\}$ is not in any clique of $\mathcal{E}_\ell$ or $\mathcal{E}_m$, the edges $\{v_j, v_m\}$ and $\{v_k, v_\ell\}$ do not appear in G. But then the vertices $v_i, v_j, v_k, v_\ell, v_m$ form a kite in G, contradicting the kite-free-ness of G.

We now consider the third case, where, without loss of generality, $v_j \notin N_{G_i}(v_i)$ and $v_k \in N_{G_i}(v_i)$. As in the second case, the edge $\{v_j, v_i\}$ appears in some clique of $\mathcal{E}_\ell$ where $\ell < i$. Since $v_k \in N_{G_i}(v_i)$, there exists a clique C of $\mathcal{E}_i$ that contains the edge $\{v_k, v_i\}$. If C does not contain $\{v_j, v_k\}$, there must exist a vertex v_p in C such that v_p is not adjacent to v_j. Otherwise, C could be expanded to include v_j, contradicting the fact that C is a maximal clique of H_i. Thus, the edges $\{v_j, v_p\}$ and $\{v_k, v_\ell\}$ do not appear in G. But then the vertices $v_i, v_j, v_k, v_\ell, v_p$ form a kite in G, again contradicting the kite-free-ness of G. □

Definition 4.16. Let G be a graph, and $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G. For each $i = 1, \ldots, n$, define

$$\begin{aligned} T_i = \{v_j : j > i,\ & v_j \text{ is adjacent to } v_i, \text{ and} \\ & \nexists\, k < i \text{ where } v_k \text{ is adjacent to both } v_i \text{ and } v_j\}. \end{aligned}$$

Define $\widehat{G}_i$ to be the subgraph of G with vertices $T_i \cup \{v_i\}$ and edges $\{\{x, v_i\} : x \in T_i\}$.

Lemma 4.17. *Let G be a graph, $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G, and $\mathcal{G}$ be an edge clique covering of G. Then the cliques of $\mathcal{G}_i$ must cover $\widehat{G}_i$.*

Proof. Let $e = \{v_i, v_j\}$ be an edge of $\widehat{G}_i$ and C a clique of $\mathcal{G}$ that covers e. Note that $i < j$. Since v_i is an endpoint of e, the least index of a vertex in C is at most i. Suppose the vertex of least index in C is v_k, where $k < i$. But then v_k is adjacent to both v_i and v_j, contradicting the construction of $\widehat{G}_i$. Thus, v_i is the vertex of least index in C, and hence C is in $\mathcal{G}_i$. Therefore, $\widehat{G}_i$ is covered by cliques of $\mathcal{G}_i$. □

Lemma 4.18. *Let G be a kite-free graph and $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G. Then the subgraphs G'_i of G generated by the Kim-Roberts elimination procedure are exactly $\widehat{G}_i$.*

Proof. Observe that only edges whose endvertices are adjacent to v_i can appear in cliques of $\mathcal{E}_i$. Let $\{v_j, v_i\}$ be an edge where $j > i$ and where there exists a $k < i$ such that v_k is adjacent to both v_i and v_j. By Lemma 4.15, $\{v_j, v_i\}$ appears in some clique of $\mathcal{E}_\ell$, $\ell \le k$, and so $\{v_j, v_i\}$ is not in G'_i. If there is no such $k < i$, then no $\mathcal{E}_\ell$ can cover the edge $\{v_j, v_i\}$ where $\ell < i$, and by Lemma 4.15, $\{v_j, v_i\}$ is then covered by a clique of $\mathcal{E}_i$. Thus, the definition of $\widehat{G}_i$ precisely describes G'_i. □

Lemma 4.19. *Let G be a kite-free graph, $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G. Then $h_G(\mathcal{E}, \mathcal{O}) = \min_{\mathcal{G}} h_G(\mathcal{G}, \mathcal{O})$, where $\mathcal{E}$ is any edge clique cover produced by the Kim-Roberts elimination procedure for the competition number on G and $\mathcal{O}$, and where the right-hand side minimum is taken over all edge clique covers $\mathcal{G}$ of G.*

Proof. Let $\mathcal{G}$ be an edge clique cover of G such that $h_G(\mathcal{G}, \mathcal{O})$ is minimized. By Lemma 4.17, $\mathcal{G}_i$ must cover $\widehat{G}_i$ for all i. But by Lemma 4.18, $G'_i = \widehat{G}_i$. Since $\mathcal{E}_i$ is chosen to be a minimum size cover of G'_i, $|\mathcal{E}_i| \le |\mathcal{G}_i|$.

We now show that $a_i^{\mathcal{E}} \ge a_i^{\mathcal{G}}$ and $b_i^{\mathcal{E}} \le b_i^{\mathcal{G}}$ for all i. Suppose, for contradiction, that there exists an i such that our desired conditions fail. Let i be the least such index. Now,

$$\begin{aligned} b_i^{\mathcal{E}} &= \max\{|\mathcal{E}_i| - a_{i-1}^{\mathcal{E}}, 0\} \\ &\leq \max\{|\mathcal{G}_i| - a_{i-1}^{\mathcal{E}}, 0\} \quad \text{since } |\mathcal{E}_i| \leq |\mathcal{G}_i| \\ &\leq \max\{|\mathcal{G}_i| - a_{i-1}^{\mathcal{G}}, 0\} \quad \text{since } a_{i-1}^{\mathcal{E}} \geq a_{i-1}^{\mathcal{G}} \text{ by assumption} \\ &= b_i^{\mathcal{G}}. \end{aligned}$$

But

$$\begin{aligned} a_i^{\mathcal{E}} &= a_{i-1}^{\mathcal{E}} - (|\mathcal{E}_i| - b_i^{\mathcal{E}}) + 1 \\ &\geq a_{i-1}^{\mathcal{E}} - (|\mathcal{G}_i| - b_i^{\mathcal{E}}) + 1 \quad \text{since } |\mathcal{E}_i| \leq |\mathcal{G}_i| \\ &\geq a_{i-1}^{\mathcal{G}} - (|\mathcal{G}_i| - b_i^{\mathcal{E}}) + 1 \quad \text{since } a_{i-1}^{\mathcal{E}} \geq a_{i-1}^{\mathcal{G}} \text{ by assumption} \\ &\geq a_{i-1}^{\mathcal{G}} - (|\mathcal{G}_i| - b_i^{\mathcal{G}}) + 1 \quad \text{from above} \\ &= a_i^{\mathcal{G}}. \end{aligned}$$

Thus $a_i^{\mathcal{E}} \geq a_i^{\mathcal{G}}$ and $b_i^{\mathcal{E}} \leq b_i^{\mathcal{G}}$ for all i. Summing over i gives $h_G(\mathcal{E}, \mathcal{O}) \leq h_G(\mathcal{G}, \mathcal{O})$. □

Theorem 4.20. *[Kim and Roberts] For a kite-free graph G, the elimination number $M(G)$ is unambiguously defined and equals the competition number $k(G)$.*

Proof. By Lemma 4.11, $k(G) = \min_{\mathcal{G}} \min_{\mathcal{O}} h_G(\mathcal{G}, \mathcal{O})$, where $\mathcal{G}$ ranges over all edge clique coverings of G and $\mathcal{O}$ ranges over all orderings of the vertices of G. By Lemma 4.10, $M(G, \mathcal{O}, \mathcal{E}) = h_G(\mathcal{E}, \mathcal{O})$ for any $\mathcal{E}$ corresponding to $\mathcal{O}$. By Lemma 4.19, $h_G(\mathcal{E}, \mathcal{O}) = h_G(\mathcal{E}', \mathcal{O})$ for any $\mathcal{E}$ and $\mathcal{E}'$ produced by the Kim-Roberts elimination procedure and corresponding to $\mathcal{O}$. It follows that $M(G, \mathcal{O}, \mathcal{E})$ and therefore the elimination number $M(G)$ is unambiguously defined. Moreover, by Lemma 4.19, $M(G) = \min_{\mathcal{O}} M(G, \mathcal{O}, \mathcal{E}) = \min_{\mathcal{O}} \min_{\mathcal{G}} h_G(\mathcal{G}, \mathcal{O})$, and therefore, $M(G) = k(G)$. □

4.4 Counterexample Showing the Kim-Roberts Elimination Procedure Does Not Always Obtain $k(G)$

Theorem 4.20 states that if G is kite-free then $M(G)$ is unambiguously defined and equals $k(G)$. Kim and Roberts asked if this is true for all graphs. To answer this question negatively, we need to demonstrate a graph G such that for each order $\mathcal{O}$ there is a choice

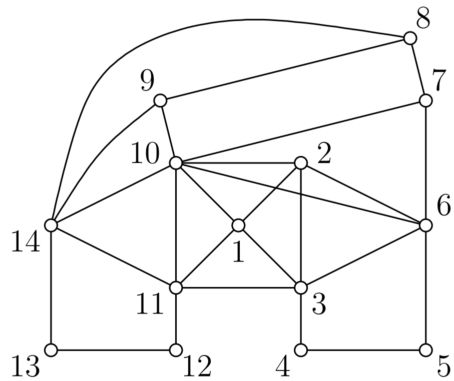

Figure 4.2: The graph L.

$\mathcal{E}$ of clique covers so that $M(G, \mathcal{O}, \mathcal{E}) > k(G)$. The graph L in Figure 4.2 is such a graph G.

From Theorem 4.20, any graph with the desired property must contain a kite. The graph L in Figure 4.2 contains two kites on the vertices $\{1, 2, 3, 10, 11\}$ and $\{1, 2, 3, 6, 10\}$. When eliminating vertices 1 or 2 first, two different clique covers of two triangles each can be used to eliminate the incident edges. One of these choices is a *good* choice for the edge clique cover, but one is a *bad* choice. Our effort in constructing the counterexample is to force 1 or 2 to be eliminated first, so that a bad choice can be made. When 1 or 2 is not eliminated first, then we will show that no choices allow the elimination procedure to attain the competition number.

Proposition 4.21. *For each ordering $\mathcal{O}$ of the vertices of the graph L in Figure 4.2, there is a choice of edge clique coverings $\mathcal{E}$ such that $M(L, \mathcal{O}, \mathcal{E}) > 2$.*

Proof. Let $\mathcal{O} = v_1, v_2, \ldots, v_{14}$ be an ordering of the vertices of L. We consider several cases:

Case 1. $v_1 = 1$.

We make the bad choice of the cliques $\{1, 2, 3\}$ and $\{1, 10, 11\}$. Any choice for v_2 other than vertex 2 can not be eliminated without increasing the number of extra vertices added to D since its remaining incident edge cannot be covered by a single clique. Thus, v_2 must be vertex 2. But after vertex 2 is eliminated, no vertex has its remaining incident edges coverable by a single clique. Thus, $M(L, \mathcal{O}, \mathcal{E}) > 2$ if $v_1 = 1$.

Case 2. $v_1 = 2$.

We make the bad choice of the cliques $\{2, 1, 3\}$ and $\{2, 6, 10\}$. Analogously to Case 1, vertex 1 is the only vertex that can then be eliminated as v_2 without increasing the number of added vertices, but after that no vertex has its remaining incident edges coverable by a single clique.

Case 3. $v_1 = 3$, 6, 10, 11, or 14.

Each of these vertices requires at least three cliques to cover its incident edges.

Case 4. $v_1 = 4$ or 5.

One of these vertices can be eliminated using two cliques, and the other is then the only vertex that can be eliminated without increasing the number of added vertices. But then no vertex has its remaining incident edges coverable by a single clique.

Case 5. $v_1 = 7$.

Vertex 7 can be eliminated with two cliques. Then vertices 8 and 9 in that order are the only vertices that can then be eliminated without increasing the number of added vertices. But after that no vertex has its remaining incident edges coverable by a single clique.

Case 6. $v_1 = 8$ or 9.

One of these vertices can be eliminated using two cliques, and then the other vertex and vertex 7 are the only vertices that can then be eliminated without increasing the number of added vertices. Vertex 7 must be eliminated after vertex 8 for this to be the case. But after that no vertex has its remaining incident edges coverable by a single clique.

Case 7. $v_1 = 12$ or 13.

One of these vertices can be eliminated using two cliques, and the other is the only vertex that can then be eliminated without increasing the number of added vertices. But after that no vertex has its remaining incident edges coverable by a single clique.

Thus, there exists a choice $\mathcal{E}$ of clique cover such that $M(L, \mathcal{O}, \mathcal{E}) > 2$ for any order $\mathcal{O}$. □

Proposition 4.22. *The competition number of the graph L in Figure 4.2 is* 2*.*

Proof. First note that there is no vertex in L whose incident edges can be covered with one clique. Thus, $k(L) \geq 2$. But the elimination procedure using the order $1, 2, \ldots, 14$ and the good choice of cliques $\{1, 2, 10\}$ and $\{1, 3, 11\}$ for vertex 1 produces an elimination number $M(L, \mathcal{O}, \mathcal{E})$ of 2. Thus, $k(L) = 2$. □

4.5 The Elimination Procedure for the Phylogeny Number

The competition number problem is essentially a problem about minimum edge clique covers, where the "value" of a cover is computed in a weighted manner. The phylogeny number problem is similar in this regard. Thus, we can formulate an elimination procedure for the phylogeny number similar to that of the competition number and obtain analogous results.

Definition 4.23. Let $D = (V, A)$ be an acyclic digraph. The *phylogeny graph* $P(D)$ is a simple undirected graph with vertex set V and with adjacencies as follows: two vertices x and y are adjacent if one of the arcs (x, y) or (y, x) is present in D, or if there exists another vertex z such that the arcs (x, z) and (y, z) are both present in D.

Definition 4.24. For a simple graph G, the phylogeny number $p(G)$ is the least number r such that there exists an acyclic digraph D on $|V(G)| + r$ vertices where G is an induced subgraph of $P(D)$.

We now give the elimination procedure for the phylogeny number. Note that the only difference from the elimination procedure for the competition number is how edges of G are "accounted for" in D.

The Elimination Procedure for the Phylogeny Number

Input: A graph G and an ordering $\mathcal{O} = v_1, v_2, \ldots, v_n$ of the vertices of G.

Output: An acyclic digraph $D := D_n$ such that G is an induced subgraph of $P(D)$.

Initialization: Set D_0 to the digraph with vertex set $V(G)$ and no arcs. D_i is an acyclic digraph constructed at the i^{th} iteration.

Set $G_1 := G$. G_i is a spanning subgraph of G that contains the edges of G that do not appear in $P(D_{i-1})$.

i^{th} Iteration, $i = 1, \ldots, n$: Set G'_i to $E_{G_i}(v_i)$, and set H_i to the subgraph of G induced by $\{v_i\} \cup \{v_j : j > i \text{ and } v_j \in N_G(v_i)\}$. Let $\mathcal{E}_i = \{C_1, \ldots, C_k\}$ be a minimum size edge covering of G'_i by maximal cliques of H_i, ordered arbitrarily. Form G_{i+1} from G_i by removing the edges of C_j from G_i for all j.

Form the digraph D_i by adding vertices and arcs to D_{i-1} as follows: Add the arcs (w, v_i) to D_i for all vertices $w \in C_1 \setminus \{v_i\}$. For each clique $C_j \in \mathcal{E}_i \setminus \{C_1\}$, add a vertex b_j to $V(D_i)$ and add the arcs (w, b_j) to D_i for each $w \in C_j$.

We also give a generalized elimination procedure for the phylogeny number.

The Generalized Elimination Procedure for the Phylogeny Number

Input: A graph G, an ordering $\mathcal{O} = v_1, v_2, \ldots, v_n$ of the vertices of G, and an edge clique covering $\mathcal{G}$ of G.

Output: An acyclic digraph $D := D_n$ such that G is an induced subgraph of $P(D)$.

Initialization: Set D_0 to the digraph with vertices $V(G)$ and no arcs. D_i is an acyclic digraph constructed at the i^{th} iteration.

i^{th} Iteration, $i = 1, \ldots, n$: Let $\mathcal{G}_i = \{C_1, \ldots, C_k\}$ be the subset of $\mathcal{G}$ where for each $C_j \in \mathcal{G}_i$, v_i is the vertex in C_j of least index. Order $\mathcal{G}_i$ arbitrarily.

Form the digraph D_i by adding vertices and arcs to D_{i-1} as follows: Add the arcs (w, v_i) to D_i for all vertices $w \in C_1 \setminus \{v_i\}$. For each clique $C_j \in \mathcal{G}_i \setminus \{C_1\}$, add a vertex b_j to $V(D_i)$ and add the arcs (w, b_j) to D_i for each $w \in C_j$.

We will first show that the generalized elimination procedure produces an acyclic digraph, and then show that for the digraph D produced by the generalized elimination procedure, $P(D)$ has G as an induced subgraph.

Lemma 4.25. *Let D be the digraph produced by the generalized elimination procedure for the phylogeny number for a graph G, a vertex ordering $\mathcal{O} = v_1, \ldots, v_n$, and an edge clique covering $\mathcal{G}$. Then all vertices in $V(D) \setminus V(G)$ are sinks, and if (v_ℓ, v_k) is an arc, then $k < \ell$. Thus, D is acyclic.*

Proof. If $b \in V(D) \setminus V(G)$, then b is a sink by construction. Now, if (v_ℓ, v_k) is an arc, then it is added to D_k in the k^{th} iteration, where v_ℓ is a vertex in C_1, a clique in $\mathcal{G}_k$. Since v_k is the vertex of least index in C_1, $k < \ell$. □

Proposition 4.26. *The generalized elimination procedure for the phylogeny number produces an acyclic digraph D such that the phylogeny graph $P(D)$ has an induced subgraph isomorphic to G.*

Proof. Let G be a graph, $\mathcal{O} = v_1, v_2, \ldots, v_n$ an ordering of the vertices of G, and $\mathcal{G}$ an edge clique covering of G. From the initialization, the vertices of G are a subset of the vertices of D. Let v_k and v_ℓ, $k < \ell$, be vertices of D that are also vertices of G. Suppose that v_k and v_ℓ are adjacent in G. Let i be the least index such that $\mathcal{G}_i$ contains a clique C that contains the edge $\{v_k, v_\ell\}$. Since $\mathcal{G}$ is an edge clique cover of G, i is well-defined. Now if $C = C_1 \in \mathcal{G}_i$, then both the arcs (v_k, v_i) and (v_ℓ, v_i) are added to D_i in the i^{th} iteration, or if $i = k$, only the arc (v_ℓ, v_k) is added. Thus, v_k and v_ℓ are adjacent in $P(D)$. Otherwise, the arcs (v_k, b_j) and (v_ℓ, b_j) are added to D_i for some b_j, and again v_k and v_ℓ are adjacent in $P(D)$.

Suppose that v_k and v_ℓ are adjacent in $P(D)$. If v_k and v_ℓ have an arc connecting them in D, then by Lemma 4.25, the arc is oriented towards v_k. Thus, in the k^{th} iteration, $v_\ell \in C_1$ for a clique $C_1 \in \mathcal{G}_k$. Since both v_k and v_ℓ are in C_1, v_k and v_ℓ must be adjacent in G. Now, if v_k and v_ℓ have incident arcs oriented towards a common vertex x, where $x \neq v_k, v_\ell$, then these arcs are added in some i^{th} iteration of the procedure. Then both v_k and v_ℓ are in the same clique $C_j \in \mathcal{G}_i$, and so must be adjacent in G. □

Note that Remark 4.4 still holds in the phylogeny number case. We also have the following lemma and proposition, whose proofs are similar to the proofs of Lemma 4.6 and Proposition 4.7.

Lemma 4.27. *Let $\mathcal{E}_i$ be the edge coverings generated by the elimination procedure for the phylogeny number for a graph G and a vertex ordering $\mathcal{O} = v_1, \ldots, v_n$. Then the set $\mathcal{E} = \bigcup_{i=1}^{n} \mathcal{E}_i$ is an edge clique covering of G.*

Proposition 4.28. *Let G be a graph and $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G. Then the number of vertices added to the digraph produced by the elimination procedure for the phylogeny number is the number of vertices added to the digraph produced by the generalized elimination procedure for the phylogeny number if the edge clique covering $\mathcal{G}$ is chosen to be $\mathcal{E}$ as defined in Lemma 4.27.*

Definition 4.29. The *phylogeny elimination number* $e_p(G, \mathcal{O}, \mathcal{G})$ of a graph G, an ordering $\mathcal{O}$ of the vertices, and an edge clique covering $\mathcal{G}$ is the number of vertices added to D so that $P(D)$ has G as an induced subgraph. Here D is the digraph produced by the generalized elimination procedure for the phylogeny number with G, $\mathcal{O}$, and $\mathcal{G}$ as inputs.

Definition 4.30. Let G be a graph, $\mathcal{O} = v_1, v_2, \ldots, v_n$ be an ordering of the vertices of G, and $\mathcal{G} = \{C_1, C_2, \ldots, C_k\}$ be an edge clique covering of G. For each vertex v_i, let $\mathcal{G}_i$ be the subset of $\mathcal{G}$ where for each $C_j \in \mathcal{G}_i$, v_i is the vertex in C_j of least index. Define

$$f_G(\mathcal{G}, \mathcal{O}) = \sum_{i=1}^{n} \max\{|\mathcal{G}_i| - 1, 0\}.$$

Lemma 4.31. *Let G be a graph, $\mathcal{O} = v_1, v_2, \ldots, v_n$ be an ordering of the vertices of G, and $\mathcal{G} = \{C_1, C_2, \ldots, C_k\}$ be an edge clique covering of G. Then $e_p(G, \mathcal{O}, \mathcal{G}) = f_G(\mathcal{G}, \mathcal{O})$.*

Proof. Note that $\mathcal{G}_i$ is defined exactly the same in both the generalized elimination procedure and in Definition 4.30. Note that in the i^{th} iteration, if $\mathcal{G}_i$ is empty, no arcs or vertices are added to D_i. If $\mathcal{G}_i$ is not empty, then $|\mathcal{G}_i| - 1$ new vertices are added as sinks to D_i. Thus, in the i^{th} iteration, $\max\{|\mathcal{G}_i| - 1, 0\}$ vertices are added to D_i, and, summing over all iterations,

$$e_p(G, \mathcal{O}, \mathcal{G}) = |V(D) \setminus V(G)| = \sum_{i=1}^{n} \max\{|\mathcal{G}_i| - 1, 0\} = f_G(\mathcal{G}, \mathcal{O}).$$

□

Lemma 4.32. *For a graph G, the phylogeny number $p(G)$ equals $\min_{\mathcal{G}} \min_{\mathcal{O}} f_G(\mathcal{G}, \mathcal{O})$, where $\mathcal{G}$ ranges over all edge clique coverings of G, and $\mathcal{O}$ ranges over all orderings of the vertices of G.*

Proof. Let $\mathcal{G} = \{C_1, C_2, \ldots, C_k\}$ be an edge clique covering of G, and $\mathcal{O} = v_1, v_2, \ldots, v_n$ be an ordering of the vertices of G. By Lemmas 4.25 and 4.26, the generalized elimination procedure produces an acyclic digraph D such that $P(D)$ has an induced subgraph isomorphic to G. By Lemma 4.31, $e_p(G, \mathcal{O}, \mathcal{G}) = f_G(\mathcal{G}, \mathcal{O})$, and so $p(G) \leq \min_{\mathcal{G}} \min_{\mathcal{O}} f_G(\mathcal{G}, \mathcal{O})$.

Now let F be an acyclic digraph that attains the phylogeny number for G; that is, $P(F)$ has an induced copy of G and $|V(F) \setminus V(G)| = p(G)$. Let $\mathcal{O} = v_1, v_2, \ldots, v_n$ be an ordering of the vertices of G such that if (v_ℓ, v_k) is an arc in F, then $k < \ell$. We construct an edge clique covering $\mathcal{G}$ of G from F as follows: For a vertex $v_i \in V(G)$, $N_F^{\text{in}}[v_i]$ induces a clique in G, and for a vertex $b \in V(F) \setminus V(G)$, $N_F^{\text{in}}(b)$ induces a clique in G. Since these are the only two ways edges can be present in G, $\mathcal{G}$ is an edge clique cover of G. Now observe that the digraph D produced by the generalized elimination procedure with $\mathcal{G}$ and $\mathcal{O}$ has the same number of vertices as F. In fact, if $C_1 \in \mathcal{G}_i$ is chosen to be the clique induced by $N_F^{\text{in}}[v_i]$, then D is isomorphic to F.

Therefore, $p(G) = |V(F) \setminus V(G)| = |V(D) \setminus V(G)| = e_p(G, \mathcal{O}, \mathcal{G}) \geq \min_{\mathcal{G}} \min_{\mathcal{O}} f_G(\mathcal{G}, \mathcal{O})$, and so $p(G) = \min_{\mathcal{G}} \min_{\mathcal{O}} f_G(\mathcal{G}, \mathcal{O})$. □

With our formula for evaluating different edge clique covers in hand, we can again turn our attention to kite-free graphs. The following lemmas are analogous to Lemmas 4.15, 4.17, and 4.18, and their proofs are the same.

Lemma 4.33. *Let G be a kite-free graph and $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G. In the elimination procedure for the phylogeny number, an edge $\{v_j, v_k\}$ with $v_j, v_k \in N_G(v_i)$ appears in some clique of $\mathcal{E}_\ell$, where $\ell \leq i$.*

Lemma 4.34. *Let G be a graph, $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G, and $\mathcal{G}$ be an edge clique covering of G. Then the cliques of $\mathcal{G}_i$ must cover $\widehat{G}_i$.*

Lemma 4.35. *Let G be a kite-free graph and $\mathcal{O} = v_1, \ldots, v_n$ be an ordering of the vertices of G. Then the subgraphs G'_i of G generated by the elimination procedure for the phylogeny* *number are exactly* $\widehat{G}_i$.

Thus we have

Lemma 4.36. *Let* G *be a kite-free graph and* $\mathcal{O} = v_1, \ldots, v_n$ *be an ordering of the vertices of* G*. Then* $f_G(\mathcal{E}, \mathcal{O}) = \min_{\mathcal{G}} f_G(\mathcal{G}, \mathcal{O})$*, where* $\mathcal{E}$ *is any edge clique cover produced by the elimination procedure for the phylogeny number on* G *and* $\mathcal{O}$*, and where the right-hand side minimum is taken over all edge clique covers* $\mathcal{G}$ *of* G*.*

Proof. Let $\mathcal{G}$ be an edge clique cover of G such that $f_G(\mathcal{G}, \mathcal{O})$ is minimized. By Lemma 4.34, $\mathcal{G}_i$ must cover $\widehat{G}_i$ for all i. But by Lemma 4.35, $G'_i = \widehat{G}_i$. Since $\mathcal{E}_i$ is chosen to be a minimum size cover of G'_i, $|\mathcal{E}_i| \leq |\mathcal{G}_i|$. Thus, $\max\{|\mathcal{E}_i| - 1, 0\} \leq \max\{|\mathcal{G}_i| - 1, 0\}$, and summing over i gives $f_G(\mathcal{E}, \mathcal{O}) \leq f_G(\mathcal{G}, \mathcal{O})$. □

We now specialize the definition of the elimination number.

Definition 4.37. Given a graph G and an ordering $\mathcal{O}$, let $\mathcal{E} = \mathcal{E}(G, \mathcal{O}) = \{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n\}$ be edge clique coverings obtained during the elimination procedure for the phylogeny number. Again, the notation is ambiguous since the way to choose the $\mathcal{E}_i$ is not completely specified in the procedure. The *phylogeny elimination number* $e_p(G)$ is the minimum of $e_p(G, \mathcal{O}, \mathcal{E})$ over all orders $\mathcal{O}$ and some $\mathcal{E}$ obtained when using $\mathcal{O}$. We will show that for certain classes of graphs, if $e_p(G)$ is this minimum and is attained for $\mathcal{O}$ and some $\mathcal{E}$, then it is attained for the same $\mathcal{O}$ and any $\mathcal{E}$ corresponding to $\mathcal{O}$. If this is the case, $e_p(G)$ is unambiguously defined.

As with $M(G)$, the determination of necessary and sufficient conditions for $e_p(G)$ to be unambiguously defined is an interesting open problem. Lemma 4.32 shows that there is always a "right" clique cover for each order such that the minimum over orders attains the phylogeny number $p(G)$. Analogous to Kim and Roberts' result, we show that for the kite-free graphs $e_p(G)$ is unambiguously defined and equals $p(G)$. We also show that the same graph L in Figure 4.2 has the property that for each order $\mathcal{O}$ there is a choice of clique cover $\mathcal{E}_i$ in the elimination procedure for the phylogeny number such that $M(G, \mathcal{O}, \mathcal{E}) > k(G)$. This demonstrates that $e_p(G)$ is not unambiguously defined and is not equal to $p(G)$ for all graphs.

Theorem 4.38. *For a kite-free graph G, the phylogeny elimination number $e_p(G)$ is unambiguously defined and equals the phylogeny number $p(G)$.*

Proof. By Lemma 4.32, $p(G) = \min_{\mathcal{G}} \min_{\mathcal{O}} f_G(\mathcal{G}, \mathcal{O})$, where $\mathcal{G}$ ranges over all edge clique coverings of G and $\mathcal{O}$ ranges over all orderings of the vertices of G. By Lemma 4.31, $e_p(G, \mathcal{O}, \mathcal{E}) = f_G(\mathcal{E}, \mathcal{O})$ for any $\mathcal{E}$ corresponding to $\mathcal{O}$. By Lemma 4.36, $f_G(\mathcal{E}, \mathcal{O}) = f_G(\mathcal{E}', \mathcal{O})$ for any $\mathcal{E}$ and $\mathcal{E}'$ produced by the elimination procedure for the phylogeny number and corresponding to $\mathcal{O}$. It follows that $e_p(G, \mathcal{O}, \mathcal{E})$ and therefore the phylogeny elimination number $e_p(G)$ is unambiguously defined. Moreover, by Lemma 4.36, $e_p(G) = \min_{\mathcal{O}} e_p(G, \mathcal{O}, \mathcal{E}) = \min_{\mathcal{O}} \min_{\mathcal{G}} f_G(\mathcal{G}, \mathcal{O})$, and therefore, $e_p(G) = p(G)$. □

Because of the similarities in the elimination procedures for the competition and phylogeny numbers, the same graph L in Figure 4.2 is also shows that the elimination procedure for the phylogeny number does not always attain $p(G)$. Both of the following propositions are proved in a fashion similar to Propositions 4.21 and 4.22 above.

Proposition 4.39. *For each ordering $\mathcal{O}$ of the vertices of the graph L in Figure 4.2, there is a choice of edge clique coverings $\mathcal{E}_i$ such that the number of added vertices by the elimination procedure for the phylogeny number is greater than* 1.

Proposition 4.40. *The phylogeny number of the graph L in Figure 4.2 is* 1.

4.6 Open Problems

Many questions still exist about the existence and efficacy of elimination procedures that calculate the competition number or the phylogeny number of a graph. Despite the existence of the graph L in Figure 4.2, the Kim-Roberts elimination procedure is still of interest, particularly in determining for which graphs the procedure calculates $k(G)$. For instance, is L the smallest graph where the elimination procedure fails, or is there a smaller example? Is there an example with only one kite? Kites without the xy edge do not always admit a choice in clique covers. Is the elimination procedure exact when there is no choice? A complete characterization of when the procedure is optimal and when it is not is still open.

Of course, all of the above questions also apply to the elimination procedure for the phylogeny number. The calculation of the competition number and the phylogeny number both are essentially problems about finding a minimum-size edge clique cover, where the "size" of the cover is computed in different ways. Because of the similarities in the problems, as well as the success of applying techniques to both problems, it seems that a reduction from one problem to another should be possible. This would eliminate the necessity of checking techniques on both problems. In fact, the phylogeny graph of an acyclic digraph D is just the competition graph of D with loops added to each vertex. However, loops are normally excluded from digraphs when considering competition graphs, and so this reduction is not very useful. No reduction is known using only acyclic digraphs. One complicating factor is that available prey are a global property in the competition number case, whereas in the phylogeny number case, the number of vertices needed to be added is strictly a local property.

The existence of the graph L where both the Kim-Roberts elimination procedure and the elimination procedure for the phylogeny number fail suggests a natural question: Can a different elimination procedure be created that succeeds for all graphs? To answer this question, a more strict definition of what constitutes an elimination procedure is needed. One reasonable condition might be to restrict what portion of the graph the procedure may consider when eliminating a vertex v. For instance, the procedure might only be able to consider vertices that are a fixed distance from v. In such instances where an elimination procedure can only consider local information, it seems unlikely that the procedure will calculate $k(G)$ or $p(G)$ for all graphs, even with the power of taking a minimum over all vertex orders. One indication supporting this view would be if it can be shown that the Kim-Roberts elimination procedure needs to solve an NP-complete problem about cliques to guarantee producing the competition number, despite the extra power of the minimum over orders. It might be possible to prove the NP-completeness using the "widgetlike" construction of the graph L. Another reasonable condition is requiring that all computation for eliminating a vertex is done in polynomial time. It seems in this case that examining factorial number of different vertex orders should give sufficient power to exactly solve either problem. However, an explicit procedure that accomplishes this is still needed.

A more general study of elimination procedures might also give insight into what graph parameters could be effectively calculated using elimination properties. Other parameters related to clique coverings are natural candidates, but perhaps other parameters such as chromatic number could also be considered.

There are many important open questions in the theory of competition graphs that this chapter has not touched on. We mention two long-standing problems here. An *interval graph* is a graph whose vertices can be associated with intervals in the real line such that two vertices are adjacent if and only if the associated intervals intersect. While investigating food webs that occur in nature, Cohen [1] empirically observed that a large number of food webs are interval graphs. If this observation is true for all food webs, the ecological implication is that there is one linear parameter (such as pH or temperature) for each ecosystem that determines the competitive relationships between species. It is not true that all competition graphs are interval graphs. However, a large literature has developed attempting to explain why competition graphs of food webs that occur in nature are often interval graphs. Of particular interest is determining which structural properties of an acyclic digraph guarantee that its competition graph is an interval graph. See [10] for a survey of these results.

A second major open problem involves a conjecture of Opsut. Opsut showed in [7] that the competition number of all line graphs is at most 2. Based on this result, Opsut conjectured that if the closed neighborhood of every vertex in a graph G can be edge-covered by at most two cliques of G, then the competition number of G is at most 2. Note that line graphs satisfy this hypothesis. Substantial progress has been made by Wang in [16], [17], and [18] towards this conjecture, but it still remains unresolved.

References

[1] J. E. Cohen, Interval Graphs and Food Webs: A Finding and a Problem, RAND Corporation Document 17696-PR, Santa Monica, CA, 1968.

[2] M. C. Golumbic, A Note on Perfect Gaussian Elimination, *J. Math. Anal. Appl.*, 64 (1978), 455-457.

[3] S.-R. Kim, The Competition Number and its Variants, in: *Quo Vadis Graph Theory?*, J. Gimbel, J. W. Kennedy, and L. V. Quintas, eds., Annals of Discrete Mathematics, 55 (1993), 313-325.

[4] S.-R. Kim and F. S. Roberts, Competition Numbers of Graphs with a Small Number of Triangles, *Discrete Applied Math.*, 78 (1997), 153-162.

[5] S.-R. Kim and F. S. Roberts, The Elimination Procedure for the Competition Number, *Ars Combinatoria,* 50 (1998), 97-113.

[6] J. R. Lundgren, Food Webs, Competition Graphs, Competition-Common Enemy Graphs, and Niche Graphs, in F. S. Roberts (ed.), *Applications of Combinatorics and Graph Theory in the Biological and Social Sciences*, vol 17 of *IMA Volumes in Mathematics and its Applications*, Springer-Verlag, New York, 1989, 221-243.

[7] R. J. Opsut, On the Computation of the Competition Number of a Graph, *SIAM J. Alg. & Discr. Meth.*, 3 (1982), 420-428.

[8] S. Parter, The Use of Linear Graphs in Gauss Elimination, *SIAM Review*, 3 (1961), 119-130.

[9] A. Raychaudhuri and F. S. Roberts, Generalized Competition Graphs and their Applications, in P. Brucker and R. Pauly (eds.), *Methods of Operations Research*, Academic Press, New York and London, 1969, 295-311.

[10] F. S. Roberts, Competition Graphs and Phylogeny Graphs, in L. Lovasz (ed.), *Graph Theory and Combinatorial Biology, Bolyai Society Mathematical Studies,* 7 (1999), *J. Bolyai Mathematical Society,* Budapest, Hungary, 333-362.

[11] F. S. Roberts, Food Webs, Competition Graphs, and the Boxicity of Ecological Phase Space, in Y. Alavi and D. Lick (eds.), *Theory and Applications of Graphs*, Springer, New York, 1978, 477-490.

[12] F. S. Roberts, *Graph Theory and its Applications to Problems of Society*, CBMS-NSF Monograph No. 29, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1978.

[13] F. S. Roberts and L. Sheng, Phylogeny Numbers, *Discrete Applied Mathematics*, 87 (1998), 213-228.

[14] F. S. Roberts and L. Sheng, Phylogeny Numbers of Graphs with Two Triangles, *Discrete Applied Math.*, 103 (2000), 191-207.

[15] D. J. Rose, Triangulated Graphs and the Elimination Process, *J. Math. Anal. Appl.*, 32 (1970), 597-607.

[16] C. Wang, *Competition Graphs, Threshold Graphs and Threshold Boolean Functions*, Ph.D. thesis, RUTCOR-Rutgers Center for Operations Research, Rutgers University, New Brunswick, NJ, 1991.

[17] C. Wang, On Critical Graphs for Opsut's Conjecture, *Ars Combinatoria*, 34 (1992), 183-203.

[18] C. Wang, Competitive Inheritance and Limitedness of Graphs, *J. Graph Theory*, 19 (1995), 353-366.

Vita

Stephen G. Hartke

1991-95 Attended St. Xavier High School, Cincinnati, Ohio. Graduated as valedictorian.

1995-99 Attended the University of Dayton, Dayton, Ohio. Graduated *summa cum laude* with a B.S. in Mathematics and Computer Science and minors in Physics and English.

1999-2004 Attended Rutgers University, Piscataway, New Jersey, for graduate work in Mathematics.

1999 (with A. W. Higgins) Maximum Degree Growth of the Iterated Line Graph, *The Electronic Journal of Combinatorics,* 6 (1999), #R28.

2000 Binary De Bruijn Cycles under Different Equivalence Relations, *Discrete Mathematics,* 215 (2000), 93-102.

2000 Summer Mathematics Research Experiences, *Proceedings of the Conference on Summer Undergraduate Mathematics Research Programs,* J. A. Gallian, ed., American Mathematical Society (2000), 267-271.

2002 (with D. Bianco, O. M. Carducci, and A. Larimer) Stable Matchings in the Couples Problem, *Morehead Electronic Journal of Applicable Mathematics*, Issue 2, MATH-2001-06, January 16, 2002.

2003 (with A. W. Higgins) Minimum Degree Growth of the Iterated Line Graph, *Ars Combinatoria*, 69 (2003), 275-283.

2003 (with M. Develin and D. P. Moulton) A General Notion of Visibility Graphs, *Discrete and Computational Geometry*, 29 (2003), no. 4, 511-524.

2003 The Elimination Procedure for the Phylogeny Number, *Ars Combinatoria*, accepted for publication.

2003 (with W. S. Diestelkamp and R. H. Kenney) On the Degree of Local Permutation Polynomials, *Journal of Combinatorial Mathematics and Combinatorial Computation*, accepted for publication.

2003 The Elimination Procedure for the Competition Number is Not Optimal, DIMACS Technical Report 2003-42.