DIMACS Technical Report 98-49
November 1998

Phylogeny Numbers for Graphs with Two Triangles

by

Fred S. Roberts[1]
Department of Mathematics, RUTCOR, and DIMACS
Rutgers University
New Brunswick, New Jersey 08903

Li Sheng[2]
Department of Mathematics and Computer Science
Drexel University, Philadelphia, PA 19014

[1]The author gratefully acknowledges the support of the National Science Foundation under grants BIR 94-12594 and SBR 97-09134 to Rutgers University.
[2]The author gratefully acknowledges the support of the National Science Foundation under grants BIR 94-12594 and SBR 97-09134 to Rutgers University.

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, Bellcore and NEC Research Institute.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

Motivated by problems of phylogenetic tree reconstruction, we introduce notions of phylogeny graph and phylogeny number. These notions are analogous to and can be considered as natural generalizations of notions of competition graph and competition number that arise from problems of ecology. Given an acyclic digraph $D = (V, A)$, define its *phylogeny graph* $G = P(D)$ by taking the same vertex set as D and, for $x \neq y$, letting $xy \in E(G)$ if and only if $(x, y) \in A$ or $(y, x) \in A$ or $(x, a), (y, a) \in A$ for some vertex $a \in V$. Given a graph $G = (V, E)$, we shall call the acyclic digraph D a *phylogeny digraph* for G if G is an induced subgraph of $P(D)$ and D has no arcs from vertices outside of G to to vertices in G. The *phylogeny number* $p(G)$ is defined to be the smallest r such that G has a phylogeny digraph D with $|V(D)| - |V(G)| = r$. In this paper, we obtain results about phylogeny number for graphs with exactly two triangles analogous to those for competition number.

1 Phylogeny Graphs and Phylogeny Numbers

The notions of competition graph and competition number were introduced by Cohen [1] and Roberts [7], respectively, to model problems in ecology, and there has since developed an extensive literature on these subjects. In Roberts and Sheng [10], we introduced related notions called phylogeny graph and phylogeny number, motivated in part by problems of phylogenetic tree reconstruction, and presented results analogous to a number of well-known results about competition graphs and competition numbers. In this paper, we calculate the phylogeny number for graphs with exactly two triangles, analogous to similar results of Kim and Roberts [4] about competition numbers.

Suppose $D = (V, A)$ is a digraph, where V is the vertex set and A the set of directed arcs. We adopt the graph-theoretical notation and terminology of Roberts [9], unless otherwise noted. The *competition graph* $K(D)$ is the undirected graph (V, E) with the same vertex set as D and with

$$xy \in E \Leftrightarrow (\exists a \in V)[(x, a) \in A \ \& \ (y, a) \in A] \tag{1}$$

In the ecological application introduced by Cohen, D represents the food web for species in the vertex set V, where an arc from u to v means that u preys on v. Then there is an edge between x and y in the competition graph if and only if x and y have a common prey. Food webs are usually acyclic, and the study of competition graphs has often been restricted to the competition graphs of acyclic digraphs. Competition graphs also arise in connection with communication over noisy channels, radio and television transmission, and models of complex economic and energy problems – see Raychaudhuri and Roberts [6] and Roberts [8] for a discussion of these applications. Roberts [7] observed that if G is any graph, then G together with sufficiently many isolated vertices is a competition graph of some acyclic digraph, and defined the *competition number* $k(G)$ to be the smallest number of such isolated vertices. Characterizing competition graphs of acyclic digraphs is thus equivalent to calculating competition numbers of arbitrary graphs. See Lundgren [5] for a survey of the extensive literature of competition graphs and competition numbers.

Given an acyclic digraph $D = (V, A)$, its *phylogeny graph* $P(D)$ is the undirected graph (V, E) with the same vertex set as D and with

$$\begin{aligned} xy \in E \Leftrightarrow (\exists a \in V)[(x, a) \in A \ \& \ (y, a) \in A] \\ \text{or } [(x, y) \in A] \text{ or } [(y, x) \in A] \end{aligned} \tag{2}$$

Given $G = (V, E)$, by adding a set I_k of k vertices, we can always build an acyclic digraph D with $V(D) = V \cup I_k$ such that for all $x, y \in V$,

$$\begin{aligned} xy \in E \Leftrightarrow (\exists a \in V \cup I_k))[(x, a) \in A(D) \ \& \ (y, a) \in A(D)] \\ \text{or } [(x, y) \in A(D)] \text{ or } [(y, x) \in A(D)] \end{aligned} \tag{3}$$

Simply add a new vertex $\alpha(x, y)$ for each edge xy in G, with arcs from x and y to $\alpha(x, y)$. Note that $P(D)$ can have edges in between some vertices of G and added vertices. However, in D we do not have arcs from added vertices to vertices in G. Any acyclic digraph D for

which Equation (3) holds and such that D has no arcs from vertices outside of G to G is called a *phylogeny digraph* for G. The *phylogeny number* $p(G)$ of G is the smallest r so that G has a phylogeny digraph D with $|V(D) - V(G)| = r$. A phylogeny digraph D for G for which $|V(D) - V(G)| = p(G)$ is called an *optimal phylogeny digraph* for G.

Some of the most helpful results in the theory of competition graphs, results that give rise to explicit formulas for the competition number and therefore are useful for making calculations and for developing and checking conjectures, are those for graphs with a small number of triangles. Results for the case of zero, one or two triangles were obtained by Kim and Roberts [4]. In [10], we obtained analogous results for phylogeny number for graphs with zero or one triangle. In this paper, we give analogous results for the case of two triangles.

The rest of this paper is organized as follows. In Section 2, we list some results about the phylogeny numbers for graphs with zero or one triangle that will be used in this paper. Section 3 deals with the phylogeny numbers for graphs with two triangles which share one of their edges. In Section 4, we calculate $p(G)$ for graphs with two triangles that are edge-disjoint.

2 Preliminaries

An *edge clique cover* of G is a collection of cliques that covers all edges of G, and a *vertex clique cover* of G is a collection of cliques that covers all vertices of G. We will denote by $\theta_e(G)$ or simply θ_e the *edge clique cover number*, i.e., the minimum number of cliques in an edge clique cover. The *vertex clique cover number* $\theta(G)$ is defined similarly. We will denote by $n(G)$ or simply n the number of vertices in G, and $e(G)$ or simply e the number edges of G. If G is a graph, $N_G(x) = N(x) = \{y : xy \in E(G)\}$ is the *open neighborhood* of vertex x in G, and $N_G[x] = N[x] = N(x) \cup \{x\}$ is the *closed neighborhood* of x. By abuse of notation, we also use $N(x)$ and $N[x]$ to denote the subgraph induced by these sets. If $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are two graphs, we will denote $G_1 \cup G_2 = (V_1 \cup V_2, E_1 \cup E_2)$. The union of two digraphs is defined accordingly.

Theorem 1 (Harary, Norman and Cartwright [3]) *Let $D = (V, A)$ be a digraph. Then D is acyclic if and only if there exists an ordering of vertices, $\sigma = [v_1, v_2, \cdots, v_n]$, such that one of the following two conditions holds,*

(1) *For all $i, j \in [1, n]$, $(v_i, v_j) \in A$ implies that $i < j$.*

(2) *For all $i, j \in [1, n]$, $(v_i, v_j) \in A$ implies that $i > j$.*

If D is a phylogeny digraph for $G = (V, E)$, we can write $D = (V \cup I_m, A)$, where I_m is the set of vertices of D not in G and, by notational convention throughout, will be assumed to have m vertices in it.

The following results from [10] will be needed here.

Lemma 2 (Roberts and Sheng [10]) *For any graph $G = (V, E)$, $p(G) \geq \theta_e - n + 1$.*

Lemma 3 (Roberts and Sheng [10]) *Given a graph $G = (V, E)$, let $G_1, G_2, \ldots, G_m$ be the connected components of G and let D_i be an optimal phylogeny digraph for G_i, $i = 1, 2, \cdots, m$. Then $D = D_1 \cup D_2 \cup \cdots \cup D_m$ is an optimal phylogeny digraph for G and $p(G) = p(G_1) + p(G_2) + \cdots + p(G_m)$.*

Given $G = (V, E)$, let $D = (V \cup I_m, A)$ be a phylogeny digraph for G and let $e = uv \in E$. We say that e is *oriented* in D if $(u, v) \in A$ or $(v, u) \in A$. We say e is *taken care of by a vertex a* in $V \cup I_m$ if $(u, a) \in A$ and $(v, a) \in A$. In the case $a \in V$, we sometimes say that *e is taken care of by* edges ua and va or say a is the *common ancestor* for the triangle with vertex set $\{u, v, a\}$, and in the case $a \notin V$, we say that e is taken care of by some added ancestor a. If $E_1 \subseteq E$, we will let $\overline{E}_1 = E - E_1$. A triangle is a *mixed triangle* relative to E_1 if some edge of the triangle is in E_1 and some edge of the triangle is in $\overline{E}_1$.

In the rest of this paper, we will intensively use the following notation. Given $G = (V, E)$, we will let E^- be the subset of E obtained by deleting all the triangle edges from E, and we will let $G^- = (V, E^-)$. Let $D = (V \cup I_m, A)$ be an optimal phylogeny digraph for $G = (V, E)$, and let $G_1 = (V_1, E_1)$ be a subgraph of G. We will denote by $D|_{G_1}$ the subgraph of D induced by vertices of G_1 and vertices from I_m that take care of some edges in E_1. I_{m_1} is the set of m_1 vertices in I_m that take care of some edges in E_1. Denote $A|_{E_1} = \{(u, v) \in A : uv \in E_1\}$, $A|_{I_{m_1}} = \{(u, a) \in A : u \in V, a \in I_{m_1}\}$, $A_1 = A|_{E_1} \cup A|_{I_{m_1}}$. Then $D|_{G_1} = (V_1 \cup I_{m_1}, A_1)$.

Lemma 4 (Roberts and Sheng [10]) *Suppose $G = (V, E)$, let $E_1 \subseteq E$ and $G_1 = (V, E_1)$. Let $D = (V \cup I_m, A)$ be an optimal phylogeny digraph for G and let $I_{\overline{m}_1}$ be the set of $\overline{m}_1$ vertices in I_m that only take care of edges in $\overline{E}_1$. If G has no mixed triangles, then $p(G) \geq p(G_1) + \overline{m}_1 \geq p(G_1) + p(\overline{G}_1)$.*

Suppose $D = (V, A)$ is a digraph and $W \subseteq V$. We say W is *shielded* (in D) if each vertex of W has at most one incoming arc in D. We shall call the orientation $D = (V, A)$ of G a *free orientation* of G if D is acyclic and V is shielded in D. A *source* for a digraph D is a vertex that has no incoming arcs in D.

Lemma 5 (Roberts and Sheng [10]) *If $G = (V, E)$ is a graph with no triangles and with k connected components, then $p(G) = e - n + k$. Moreover, if x is an arbitrary vertex of G, there is an optimal phylogeny digraph for G with x a source vertex and with V shielded.*

The results of [10] that we extend to the case of two triangles are the following.

Theorem 6 (Roberts and Sheng [10]) *If $G = (V, E)$ is a connected graph with no triangles, then $p(G) = e - n + 1$.*

Theorem 7 (Roberts and Sheng [10]) *Let $G = (V, E)$ be a connected graph with exactly one triangle. Then*

$$p(G) = \begin{cases} e - n & \text{if } G^- \text{ has three components} \\ e - n - 1 & \text{if } G^- \text{ has one or two components.} \end{cases}$$

3 Phylogeny Number for Graphs with Two Triangles that Share One of their Edges

Assume that the connected graph $G = (V, E)$ has exactly two triangles. In this section, we consider the case when these two triangle share one of their edges. In this case, we denote the vertex set for the two triangles by x, u, v and y, u, v respectively, with uv being their common edge. It is easy to see that G^- has at most four components.

We say a graph $G = (V, E)$ is *separable* by a vertex $w \in V$ into two subgraphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ if $V = V_1 \cup V_2$ and $E = E_1 \cup E_2$ are such that $V_1 \cap V_2 = \{w\}$.

It follows that if $G = (V, E)$ is separable by a vertex $x \in V$ into two subgraphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, then $E_1 \cap E_2 = \emptyset$, and for vertices $u \in V_1 - \{w\}, v \in V_2 - \{w\}$, $uv \notin E$.

Lemma 8 *Let graph $G = (V, E)$ be separable by vertex w into two graphs $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, and let $D = (V \cup I_m, A)$ be an optimal phylogeny digraph for G. Then $D|_{G_1} = (V_1 \cup I_{m_1}, A_1)$, $D|_{G_2} = (V_2 \cup I_{m_2}, A_2)$ are phylogeny digraphs for G_1 and G_2 respectively, and $p(G) \geq p(G_1) + p(G_2)$.*

Proof. It is routine to check by definition that $D|_{G_1}$ and $D|_{G_2}$ are phylogeny digraphs for G_1 and G_2 respectively. So $p(G) = m = m_1 + m_2 \geq p(G_1) + p(G_2)$. □

Lemma 9 *Let graph $G = (V, E)$ be separable by vertex w into two graphs G_1 and G_2. If at least one of G_1 and G_2 has an optimal phylogeny digraph with no incoming arcs towards w, then $p(G) = p(G_1) + p(G_2)$.*

Proof. It follows from Lemma 8 that $p(G) \geq p(G_1) + p(G_2)$. Conversely, to show that $p(G) \leq p(G_1) + p(G_2)$, let $D_1 = (V_1 \cup I_{m_1}, A_1)$, $D_2 = (V_2 \cup I_{m_2}, A_2)$ be optimal phylogeny digraphs for G_1, G_2, respectively, with $I_{m_1} \cap V = I_{m_2} \cap V = \emptyset$. Assume without loss of generality that there is no incoming arcs toward w in D_2. It is easy to see that $I_{m_1} \cap I_{m_2} = \emptyset$. Let $m = m_1 + m_2$, $I_m = I_{m_1} \cup I_{m_2}$, and let $D = D_1 \cup D_2 = (V \cup I_m, A_1 \cup A_2)$. We will show that D is a phylogeny digraph for G. It is easy to see that $uv \in E = E_1 \cup E_2$ implies that uv is oriented in D or $(u, a), (v, a) \in A$ for some $a \in V \cup I_m$. Conversely, it is also easy to see that if uv is oriented in D or if uv is taken care of by some added vertex in I_m, then $uv \in E_1 \cup E_2 = E$. Assume now that $(u, a), (v, a) \in A(D)$ for some $a \in V$. Then $ua, va \in E_1 \cup E_2 = E$. If $a = w$, then $(u, a), (v, a) \in A_1$ since D_2 has no incoming arcs toward $w = a$. If $a \in V_1 - \{w\}$, then $u, v \notin V_2 - \{w\}$ since $ua, va \in E$. So, $u, v \in V_1$, which implies that $(u, a), (v, a) \in A_1$. Thus, we always have $uv \in E_1 \subseteq E$ since D_1 is a phylogeny digraph for G_1. The proof is similar if $a \in V_2 - \{w\}$ □

Lemma 10 *Let graph $G = (V, E)$ be separable by vertex w into two graphs G_1 and G_2. If at least one of G_1 and G_2 is triangle free, then $p(G) = p(G_1) + p(G_2)$.*

Proof. Without loss of generality, assume that G_2 is triangle free. By Lemma 5, G_2 has an optimal phylogeny digraph with w being a source vertex. The result follows directly from Lemma 9. □

For a vertex $w \in V$, we will denote by G_w the component of G^- that contains vertex w, and we denote $n_w = n(G_w)$, $e_w = e(G_w)$.

Theorem 11 *Let $G = (V, E)$ be a connected graph with exactly two triangles which share one of their edges. If G^- has four components, then $p(G) = e - n - 1$.*

Proof. Let vertices for these two triangles be x, u, v, y with the edge uv being their common edge. Let G_0 be the subgraph of G induced by the four triangle vertices x, y, u, v. It is easy to see that $p(G_0) = 0$. By Theorem 6, $p(G_w) = e_w - n_w + 1$ for vertices $w = x, y, u$ or v. It follows from Lemma 10 that $p(G) = p(G_0) + p(G_x) + p(G_y) + p(G_u) + p(G_v) = (e_x - n_x + 1) + (e_y - n_y + 1) + (e_u - n_u + 1) + (e_v - n_v + 1) = (e - 5) - n + 4 = e - n - 1$. □

Theorem 12 *Let $G = (V, E)$ be a connected graph with exactly two triangles which share one of their edges. Then, $p(G) \geq e - n - 2$. Moreover, $p(G) = e - n - 2$ if G^- has one or two components.*

Proof. Let vertices for these two triangles be x, u, v, y with the edge uv being their common edge. Notice that $\theta_e = (e-5)+2 = e-3$. It follows from Lemma 2 that $p(G) \geq (e-3)-n+1 = e - n - 2$. We shall show now that $p(G) \leq e - n - 2$ when G^- has one or two components.

Suppose G^- has one components. Then G^- is a connected graph with no triangles, and by Theorem 6, $p(G^-) = (e - 5) - n + 1 = e - n - 4$. Let D^- be an optimal phylogeny digraph for G^-, let D be such that $V(D) = V(D^-) \cup \{a, b\}$ and $A(D) = A(D^-) \cup \{(u, a), (v, a), (x, a), (u, b), (v, b), (y, b)\}$. Then D is a phylogeny digraph for G with $|V(D)| - |V(G)| = |V(D^-)| + 2 - |V(G^-)| = p(G^-) + 2 = e - n - 2$, and so $p(G) \leq e - n - 2$.

Assume now that G^- has two components C_1 and C_2.

Let $E' = E - \{xu, xv\}$. Then without loss of generality we can assume that $G' = (V, E')$ is connected. Now G' is a connected graph with exactly one triangle, and notice that if we delete all three triangle edges in G', the resulting graph is G^-, which has two components. By Theorem 7, $p(G') = (e-2)-n-1 = e-n-3$. Let $D' = (V', A')$ be an optimal phylogeny digraph for G' and let D be such that $V(D) = V' \cup \{a\}$ and $A(D) = A' \cup \{(x, a), (u, a), (v, a)\}$. Then D is a phylogeny digraph for G with $|V(D)| - |V(G)| = |V(D')| + 1 - |V(G')| = p(G') + 1 = e - n - 2$, and so $p(G) \leq e - n - 2$. □

We define $\omega(G)$ to be the number of vertices in a maximum clique of G.

Lemma 13 *Let $G = (V, E)$ have $\omega(G) \leq 3$, let $D = (V \cup I_m, A)$ be an optimal phylogeny digraph for G. Let $E_1 \subseteq E$ and let $G_1 = (V, E_1)$. Assume that every mixed triangle in G consists of one edge of E_1 and two edges of $\overline{E}_1$. Suppose that for any $e_1 \in E_1$ which is an edge for a mixed triangle, e_1 is either oriented in D or is taken care of by an added ancestor. Then $p(G) \geq p(G_1) + \overline{m}_1$.*

Proof. Let $I_{\overline{m}_1}$ be the set of $\overline{m}_1$ vertices in I_m that only take care of edges in $\overline{E}_1$. Note that for any $a \in I_{m_1} = I_m - I_{\overline{m}_1}$, a takes care of some edge of E_1 and a may also take care of some edge in $\overline{E}_1$. Let $A|_{E_1} = \{(u,v) \in A : uv \in E_1\}$, $A|_{I_{m_1}} = \{(u,a) \in A : a \in I_{m_1}$ takes care of some edge $uv \in E_1\}$, and let $D|_{G_1} = (V_1, A_1) = (V \cup I_{m_1}, A|_{E_1} \cup A|_{I_{m_1}})$. Note that unlike $A|_{I_{m_1}}$ in Lemma 4, $(u,a) \in A$ for vertices $u \in V$ and $a \in I_{m_1}$ doesn't mean that $(u,a) \in A|_{I_{m_1}}$. Notice also that I_{m_1} and $I_{\overline{m}_1}$ partition I_m, and so $p(G) = m = m_1 + \overline{m}_1$. It is easy to see that $(u,v) \in A_1$ and $u, v \in V$ implies that $(u,v) \in A|_{E_1}$, which in turn implies that $uv \in E_1$. We will show below that $D|_{G_1}$ is a phylogeny digraph for G_1, that is for any $u, v \in V$, we will show that $uv \in E_1$ if and only if $(u,v) \in A_1$ or $(v,u) \in A_1$ or $(u,a), (v,a) \in A_1$ for some $a \in V \cup I_{m_1}$. Then it follows that $m_1 \geq p(G_1)$, which implies that $p(G) = m = m_1 + \overline{m}_1 \geq p(G_1) + \overline{m}_1$.

$''\Rightarrow''$. Assume that $uv \in E_1$. If uv is oriented in D, then (u,v) *or* $(v,u) \in A|_{E_1} \subseteq A_1$. If $uv \in E_1$ is taken care of by a vertex $a \in I_m$, then $a \in I_{m_1}$ and $(u,a), (v,a) \in A|_{I_{m_1}} \subseteq A_1$. Now assume that uv is not oriented in D and uv is not taken care of by a vertex $a \in I_m$. Then, we must have that $(u,a), (v,a) \in A$ for some $a \in V$. Then $ua \in E$ and $va \in E$. We must have that $ua, va \in E_1$, for otherwise, uv would be an edge for a mixed triangle and thus be either oriented in D or be taken care of by some added ancestor, a contradiction. It follows that $(u,a), (v,a) \in A_1$.

$''\Leftarrow''$. Conversely, if (u,v) *or* $(v,u) \in A_1$, then $uv \in E_1$. If $(u,a), (v,a) \in A_1$ for some $a \in V$, then $ua, va \in E_1$ and so $uv \in E_1$ for otherwise G would have a mixed triangle with one edge in $\overline{E}_1$ and two edges in E_1. Assume now that $(u,a), (v,a) \in A_1$ for some $a \in I_{m_1}$. Then $uv \in E$. If on the contrary that $uv \notin E_1$, then a will have to take care of some edges $ux \in E_1$ and $vy \in E_1$ for some $x, y \in V$. Note that $x \neq y$ for otherwise, u, v, x would form a mixed triangle with two edges of E_1 and one edge of $\overline{E}_1$. But then, u, v, x, y would be a clique of G of 4 vertices since u, v, x, y have a common ancestor a, which contradicts that $\omega(G) \leq 3$. □

Theorem 14 *Let $G = (V, E)$ be a connected graph with exactly two triangles which share one of their edges. Let vertices for these two triangles be x, u, v, y with the edge uv being their common edge. If G^- has three components with $G_x = G_y$, then $p(G) = e - n - 1$.*

Proof. To show $p(G) \leq e - n - 1$, let $G' = (V, E')$ where $E' = E - \{uv, uy, vy\}$. Then since G' is connected with no triangles, $p(G') = (e-3) - n + 1 = e - n - 2$ by Theorem 6. Let $D' = (V', A')$ be an optimal phylogeny digraph for G'. Consider a digraph D obtained from D' by adding a vertex a to I_{e-n-2} and adding three arcs $(u,a), (v,a), (y,a)$ to A'. It is easy to see that D is a phylogeny digraph for G with $|V(D| - |V(G)| = |V(D')| + 1 - |V(G')| = p(G') + 1 = e - n - 1$, so $p(G) \leq e - n - 1$.

We will now show $p(G) \geq e - n - 1$ and the conclusion follows immediately. Suppose first that there are two added ancestors a, b for the two triangles. Then let $G_1 = G^-$. G_1 is a graph with three components and with no triangles, so by Lemma 5, $p(G_1) = (e-5) - n + 3 = e - n - 2$. Notice that $\overline{E}_1 = \{xu, xv, yu, yv, uv\}$, so $a, b \in I_{\overline{m}_1}$, which implies that $\overline{m}_1 \geq 2$. Since there is no mixed triangles in G, by Lemma 4, $p(G) \geq p(G_1) + \overline{m}_1 \geq e - n - 2 + 2 = e - n$, which contradicts that $p(G) \leq e - n - 1$.

So there is at most one added ancestor for edges in the two triangles.

Case 1. There is one added ancestor a for edges in the two triangles. Then one of the two triangles must have at least two oriented edges. So we can assume without loss of generality that xv is oriented and a takes care of some edge of the triangle y, u, v. Let $E_1 = E - \{xu, uv, yu, yv\}$. $G_1 = (V, E_1)$ has no triangles and G_1 has two components, by Lemma 5, $p(G_1) = (e-4) - n + 2 = e - n - 2$. Note that $a \in I_{\overline{m}_1}$, which implies that $\overline{m}_1 \geq 1$. By Lemma 13, $p(G) \geq p(G_1) + \overline{m}_1 \geq e - n - 2 + 1 = e - n - 1$.

Case 2. There is no added ancestor for edges in the two triangles. Then each of the two triangles must have at least two oriented edges which take care of the third edge in that triangle. So we can assume without loss of generality that both xv and yv are oriented. Let $E_1 = E - \{ux, uv, uy\}$. Then $G_1 = (V, E_1)$ has no triangles with two components, so by Lemma 5, $p(G_1) = (e-3) - n + 2 = e - n - 1$. By Lemma 13, $p(G) \geq p(G_1) + \overline{m}_1 \geq p(G_1) = e - n - 1$.

□

Lemma 15 *Let $T = (V, E)$ be a tree and let $r \in V$. Let $D' = (V', A')$ be a free orientation for a subtree T' with r a source, and let x be a vertex that is connected to r by a path containing no vertices of $V' - \{r\}$. Then D' can be extended to be a free orientation for the tree T with x a source vertex.*

Proof. We will extend D' to $D = (V, A)$ as follows. Since T is a tree, there is one and only one path, $P_{xr} : x = x_0, x_1, \cdots, x_t = r$, connecting x with r. We first orient the edge $x_i x_{i+1}$ on this path from x_i to x_{i+1}, which is valid because no such edge has been oriented in A'. Now for any other edge $e = uv$ that has not been oriented, there is a unique path starting at x that has e as its last edge; we orient the edge uv by (u, v) if the direction the edge e is traversed on this path is from u to v. It is easy to see that x is a source vertex of D. It remains to prove that D is a free orientation for T. If $(u, w) \in A - A'$, then there exists a path P_{xu} in T from x to u with w not on the path. If $(u, w) \in A'$, then there is a path P_{ru} in T' from r to u with w not on the path since D' is a free orientation rooted at r. Therefore, together with the path P_{xr}, we again have a path in T from x to u with w not on the path. Now suppose that $(u, w), (v, w) \in A$ for some vertices $u, v, w \in V$. Then there are paths P_{xu}, P_{xv} in T from x to u, v respectively with w not on either of the paths, and these two paths in T together with edges uw and vw contain a cycle in T, a contradiction. It follows immediately that D is a free orientation of T with x a source vertex. □

Lemma 16 *Let $G = (V, E)$ be a connected graph with no triangles. Let $D' = (V', A')$ be a free orientation for some tree T' in G with r a source vertex, and let x be a vertex that is connected to r by a path containing no vertices of $V' - \{r\}$. Then D' can be extended to be an optimal phylogeny digraph for G with x a source vertex and with V shielded.*

Proof. Let T' be a subtree of some spanning tree T of G. By Lemma 15, D' can be extended to be a free orientation, say $D_T = (V, A_T)$, for T with x a source vertex. Notice that $p = e-n+1$ is the number of edges of G not in T. Let $D_G = (V \cup I_p, A)$ be the digraph obtained from D_T by adding a vertex α_{xy}, and two arcs $(x, \alpha_{xy}), (y, \alpha_{xy})$, for each edge xy not in T. It is easy to see that D is a phylogeny digraph for G with x a source vertex and V shielded. Since $p(G) = e - n + 1$ by Theorem 6, it follows that D_G is optimal. □

Theorem 17 *Let $G = (V, E)$ be a connected graph with exactly two triangles which share one of their edges. Let vertices for these two triangles be x, u, v, y with the edge uv being their common edge. Then*

$$p(G) = \begin{cases} e-n-1 & \text{if } G^- \text{ has four components,} \\ & \text{or if } G^- \text{ has three components with } G_x = G_y \\ e-n-2 & \text{otherwise} \end{cases}$$

Proof. Note that G^- can have at most four components. If G^- has one or two components, then $p(G) = e - n - 2$ by Theorem 12. If G^- has four components or if G^- has three components with $G_x = G_y$, then $p(G) = e-n-1$ by Theorem 11 and Theorem 14. Assume now that G^- has three components and $G_x \neq G_y$. Note that $p(G) \geq e-n-2$ by Theorem 12. We show that $p(G) \leq e - n - 2$. By symmetry, we only need to consider the following two cases.

Case 1. $G_x = G_v$. Let $G' = (V, E')$ with $E' = E - \{xu, xv, yv\}$ and let D' be an optimal phylogeny digraph for G'. Then G' is a connected graph with no triangles. By Theorem 6, $p(G') = (e-3) - n + 1 = e - n - 2$ and by Lemma 16, we can assume without loss of generality that $(v, u), (u, y) \in A(D')$ and u and y each has no other incoming arcs in D'. Let D be a digraph obtained from D' by adding $(x, u), (v, y)$ to A'. It is routine to see that D is acyclic, and that D is a phylogeny digraph for G with $|V(D)| - |V(G)| = |V(D')| - |V(G')| = p(G') = e - n - 2$, so $p(G) \leq e - n - 2$.

Case 2. $G_u = G_v$. Analogously, let $G' = (V, E')$ with $E' = E - \{uv, xv, yv\}$. Let $D' = (V', A')$ be an optimal phylogeny digraph for G'. Then $G' = (V, E')$ is a connected graph with no triangles. Again by Theorem 6, $p(G') = (e-3) - n + 1 = e - n - 2$ and by Lemma 16, we can assume without loss of generality that $(u, x), (u, y) \in A'$, and x, y have no other incoming arcs in D'. Let D be a digraph obtained from D' by adding $(v, x), (v, y)$ to A'. Then it is easy to see that D is acyclic, and that D is a phylogeny digraph for G, so $p(G) \leq p(G') = e - n - 2$.

□

Theorem 17 should be compared to the analogous result of Kim and Roberts [4], which says that if G is a connected graph with exactly two triangles and they share one of their edges, then $k(G) = e - n - 1$ or $k(G) = e - n$.

4 Phylogeny Number for Graphs with Two Triangles that are edge-disjoint

We will now consider the case when the connected graph $G = (V, E)$ has exactly two triangles, and the two triangles share none of their edges. We will denote the vertices for the two triangles by x, y, z and u, v, w. Then by the connectivity of G, we will assume that vertex z and w are connected by some path in G and it follows that G^- can have at most five components. We call this the *canonical notation for the triangles.* Again, let G_x, G_y, G_u, G_v be the components of G^- that contain x, y, u, v respectively, and let G_{zw} be the component of G^- that contains z and w. Note also that a special case is that vertex z and w coincide. For $S \subseteq V$, let $G(S)$ be the subgraph of G induced by S.

Theorem 18 *Let $G = (V, E)$ be a connected graph with exactly two triangles that are edge-disjoint. If G^- has five components, then $p(G) = e - n - 1$.*

Proof. Use the canonical notation for the triangles. Let $G_0 = G(\{x, y, z\}) \cup G(\{u, v, w\})$. Then $p(G_0) = 0$. By Theorem 6, $p(G_{zw}) = e_{zw} - n_{zw} + 1$ and $p(G_t) = e_t - n_t + 1$ for vertices $t = x, y, u$ or v. It follows from Lemma 10 that $p(G) = p(G_0) + p(G_x) + p(G_y) + p(G_u) + p(G_v) + p(G_{zw}) = (e_x - n_x + 1) + (e_y - n_y + 1) + (e_u - n_u + 1) + (e_v - n_v + 1) + (e_{zw} - n_{zw} + 1) = (e - 6) - n + 5 = e - n - 1$. □

Lemma 19 *Let $G = (V, E)$ have exactly one triangle. If G^- has three components, then $p(G) = e - n$. Moreover, for any vertex x of the triangle, G has an optimal phylogeny digraph $D = (V \cup I_{e-n}, A)$ with x a source vertex.*

Proof. By Theorem 7, $p(G) = e - n$. Let x, y, z be the three vertices of the triangle. Let G_x, G_y and G_z be the three components of G^-. By Lemma 5, we can assume that D_x, D_y and D_z are optimal phylogeny digraphs for G_x, G_y and G_z, with x, y, z being their sources vertices respectively. Then the digraph $D = D_x \cup D_y \cup D_z \cup \{(x, z), (y, z)\}$ is a phylogeny digraph for G with x a source vertex. Note that $V(D_t) - V(G_t), t = x, y, z$, are pairwise disjoint. Since $|V(D)| - |V(G)| = \sum_{t \in \{x,y,z\}} (|V(D_t)| - |V(G_t)|) = \sum_{t \in \{x,y,z\}} (e_t - n_t + 1) = (e - 3) - n + 3 = e - n$, it follows that D is optimal. □

Theorem 20 *Let $G = (V, E)$ be a connected graph with exactly two triangles that are edge-disjoint. Then, $p(G) \geq e - n - 3$. Moreover, $p(G) = e - n - 3$ if G^- has one or two components.*

Proof. Notice that $\theta_e = (e-6)+2 = e-4$. It follows from Lemma 2 that $p(G) \geq \theta_e - n + 1 = (e - 4) - n + 1 = e - n - 3$. We shall show now that $p(G) \leq e - n - 3$ when G^- has one or two components.

Use the canonical notation for triangles. Suppose G^- has one components. If $G' = G^-$, then G' is a connected graph with no triangles. By Theorem 6, $p(G') = (e - 6) - n + 1 = e-n-5$. Let $D' = (V', A')$ be an optimal phylogeny digraph for G' and let D be obtained from

D' by adding vertices a and b to V' and adding arcs $(x,a),(y,a),(z,a),(u,b),(v,b),(w,b)$ to A'. It is easy to see that D is a a phylogeny digraph for G with $|V(D)| - |V(G)| = |V(D')| + 2 - |V(G')| = p(G') + 2 = e - n - 3$, and so $p(G) \leq e - n - 3$.

Assume now that G^- has two components and denote the vertex sets for the two components by C_1 and C_2 respectively. Let $E_1 = E - \{xy, xz, yz\}$ and $E_2 = E - \{uv, uw, vw\}$. We will show first that either $G_1 = (V, E_1)$ or $G_2 = (V, E_2)$ is connected. For if G_1 is not connected, then either x or y must be in a component different from the component of G^- containing u, v or w. But since there are only two components, it follows that u, v, w must be in the same component of G^- and so G_2 is connected.

So, without loss of generality we can assume that G_1 is connected and let $G' = G_1$. Then G' is a connected graph with exactly one triangle. Notice also that if we delete all the three triangle edges of G', the resulting graph will be G^-, which has two components. By Theorem 7, $p(G') = (e-3) - n - 1$. Let $D' = (V', A')$ be an optimal phylogeny digraph for G', let D be obtained from D' by adding a vertex a to V' and adding arcs $(x,a),(y,a),(z,a)$ to A'. It is routine to check that D is an phylogeny digraph for G with $|V(D)| - |V(G)| = |V(D')| + 1 - |V(G')| = p(G') + 1 = e - n - 3$, and so $p(G) \leq e - n - 3$. □

Lemma 21 *Let $G = (V, E)$ be a connected graph with exactly one triangle and let D_u be a free orientation on a tree of G^- with some vertex u a source vertex for D_u. If there is a triangle vertex connected to u by a path in G^- that contains no vertices of $V(D_u) - \{u\}$, then D_u can be extended to be an optimal phylogeny digraph for G with the nontriangle vertices shielded.*

Proof. Let vertices x, y, z be the three triangle vertices. We consider the following three cases.

Case 1 Suppose G^- has three components C_1, C_2 and C_3 with n_i vertices and e_i edges in component C_i. We can assume without loss of generality that $x \in C_1, y \in C_2, z \in C_3$ and $u \in C_3$. Then, by Lemma 16, D_u can be extended to be an optimal phylogeny digraph D_3 for C_3 with z a source vertex and with C_3 shielded. By Lemma 5, there are optimal phylogeny digraphs D_i for $C_i, i = 1, 2$, with C_i shielded. Thus by Lemma 3, $D^- = D_1 \cup D_2 \cup D_3$ is an optimal phylogeny digraph for G^- with z a source vertex and V shielded, and by Lemma 5 $p(G^-) = (e-3) - n + 3 = e - n$. Now let D be a digraph obtained from D^- by adding the arcs $(x,z),(y,z)$ to $A(D^-)$, it is easy to see that D is a phylogeny digraph for G with $V - \{z\}$ shielded. Now since $|V(D)| - |V(G)| = |V(D^-)| - |V(G)| = p(G^-) = e - n$, it follows from Theorem 7 that D is optimal.

Case 2 Assume now that G^- has two components C_1, C_2 and let $x, y \in C_1, z \in C_2$. If $u \in C_2$, then by Lemma 16, D_u can be extended to be an optimal phylogeny digraph D_2 for C_2 with z a source vertex and C_2 shielded. By Lemma 5, there is an optimal phylogeny digraph D_1 for C_1 with C_1 shielded. By Lemma 3, $D^- = D_1 \cup D_2$ is an optimal phylogeny digraph for G^- with z a source vertex and V shielded, and by

Lemma 5, $p(G^-) = (e-3) - n + 2 = e - n - 1$. If $u \in C_1$, then either x or y, without loss of generality x is connected to u by a path that contains no vertices of $V(D_u) - \{u\}$. By Lemma 16, D_u can be extended to be an optimal phylogeny digraph D_1 for C_1 with x a source vertex and C_1 shielded. By Lemma 5, there is an optimal phylogeny digraph D_2 for C_2 with z a source vertex and C_2 shielded. Thus, as in the case of $u \in C_2$, $D^- = D_1 \cup D_2$ is an optimal phylogeny digraph for G^- with z a source vertex and V shielded, and $p(G^-) = (e-3) - n + 2 = e - n - 1$. In both situations $u \in C_2$ and $u \in C_1$, we obtain a digraph D from D^- the same way as in case (1), with $|V(D)| - |V(G)| = p(G^-) = e - n - 1$. It follows from Theorem 7 that D is an optimal phylogeny digraph for G with $V - \{z\}$ shielded.

Case 3 Finally, assume that G^- has one component. By Lemma 16, we can extend D_u to be an optimal phylogeny digraph D^- for G^- with V shielded. Let D be the digraph obtained from D^- by adding a vertex a to $V(D^-)$ and adding the arcs $(x,a),(y,a),(z,a)$ to $A(D^-)$. It is easy to see that D is still acyclic and that D is a phylogeny digraph for G with V shielded. Since G^- is a connected graph with no triangles, by Lemma 5, $p(G^-) = (e-3)-n+1 = e-n-2$. Therefore $|V(D)|-|V(G)| = |V(D^-)|+1-|V(G^-)| = p(G^-)+1 = e-n-1$, and it follows from Theorem 7 that D is optimal.

□

Theorem 22 *Let $G = (V,E)$ be a connected graph with exactly two triangles that are edge-disjoint. If G^- has four components, then $p(G) = e - n - 2$.*

Proof. Use the canonical notation for the triangles. Since G^- has four components, it must be the case that two of the five components $G_x, G_y, G_{zw}, G_u, G_v$ are the same, and the remaining three are different. By symmetry, we just need to consider the following two cases.

Case 1. $G_u = G_v$ or $G_u = G_{zw}$.

It is easy to see that in this case G is separable by z into two subgraphs G_1 and G_2, where G_1 contains exactly one triangle x,y,z and G_2 contains exactly one triangle u,v,w. Notice that G_1^-, the graph obtained from G_1 by deleting its three triangle edges, has three components, and similarly G_2^- has two components. By Theorem 7, $p(G_1) = e_1 - n_1$, $p(G_2) = e_2 - n_2 - 1$. By Lemma 19, G_1 has an optimal phylogeny digraph that has no incoming arcs toward z. By Lemma 9, $p(G) = p(G_1) + p(G_2) = (e_1+e_2) - (n_1+n_2) - 1 = e - (n+1) - 1 = e - n - 2$.

Case 2. $G_x = G_u$.

To show $p(G) \leq e - n - 2$, let $E' = E - \{uv\}$ and let $G' = (V,E')$. Then G' is a connected graph with one triangle. Note that G'^-, the graph obtained from G' by deleting its three triangle edges, has two components. It follows from Theorem 7 that $p(G') = (e-1) - n - 1 = e - n - 2$. Now let $D' = (V',A')$ be an optimal phylogeny digraph for G'. By Lemma 21, we can assume that $(w,v) \in A'$ and v has no other

incoming arcs in D'. Let D be the digraph obtained from D' by taking the same vertex set as D' and adding arc (u,v) to A'. It is easy to see that D is acyclic and it is a phylogeny digraph for G with $|V(D)|-|V(G)| = |V(D')|-|V(G')| = p(G') = e-n-2$, and so $p(G) \leq e-n-2$. We will show now that $p(G) \geq e-n-2$ and the Theorem follows immediately. Notice first that in any optimal phylogeny digraph D for G, there is no added ancestor for any triangle edge, for otherwise, we can assume that there is an added ancestor a for some edge in the triangle x,y,z. Let $G_1 = G^-$. Then $\overline{E}_1$ is the set of triangle edges of G. Let $I_{\overline{m}_1}$ be the set of $\overline{m}_1$ vertices in I_m that only take care of edges in $\overline{E}_1$. Then $a \in I_{\overline{m}_1}$, and so, $\overline{m}_1 \geq 1$. Notice that G_1 is a graph with four components and with no triangles, so $p(G_1) = (e-6)-n+4 = e-n-2$ by Lemma 5. Since there is no mixed triangle of G, by Lemma 4, $p(G) \geq p(G_1) + \overline{m}_1 \geq e-n-1$, a contradiction. So, each of the two triangles must have at least two oriented edges which take care of the remaining edge. If yz is oriented, then let $E_1 = E - \{yx, zx\}$. Otherwise, we must have that xy is oriented, and so let $E_1 = E - \{xz, yz\}$. Then $G_1 = (V, E_1)$ is a connected graph with one triangle. It is easy to see that G_1^-, the graph obtained from G_1 by deleting all three of its triangle edges, has three components. By Theorem 7, $p(G_1) = (e-2)-n = e-n-2$. Thus, by Lemma 13, $p(G) \geq p(G_1) + \overline{m}_1 \geq p(G_1) = e-n-2$.

□

Lemma 23 *Let $G = (V,E)$ have exactly one triangle with x,y,z being the vertices for the triangle. If G^- has two components C_1, C_2 with $x,y \in C_1$ and $z \in C_2$, then $p(G) = e-n-1$. Let D be an optimal phylogeny digraph for G and let $w \in C_2 - \{z\}$. Then*

(1) *$(x,z) \in A(D)$ and $(y,z) \in A(D)$.*

(2) *$(z_1, w) \in A(D)$ for some $z_1 \in C_2$.*

Proof. By Theorem 7, $p(G) = e-n-1$. Again, by abuse of notation, let $C_i, i = 1,2$, represent the vertex set of the component as well as the component itself.

(1) Let $I_m = V(D) - V(G)$ with $m = e-n-1$ and let $E_1 = E - \{xz, yz\}$, $G_1 = (V, E_1)$. Then G_1 is a graph with no triangles and with two components. By Lemma 5, $p(G_1) = (e-2)-n+2 = e-n$. Notice that G has only one mixed triangle with one edge of E_1 and two edges of $\overline{E}_1 = \{xy, xz\}$. If $(x,z) \notin A$ or $(y,z) \notin A$, then $e_1 = xy$ can't be taken care of by z and so it has to be either oriented in D or taken care of by some added ancestor. By Lemma 13, $p(G) \geq p(G_1) = e-n$, a contradiction.

(2) If not, then there is a path $P_1 : z = z_0, z_1, \cdots, z_k = w$ in C_2 connecting z and w. If all edges in P_1 are oriented in D, then $(z, z_1) \in A(D)$, for otherwise $\{z_1, x, y\}$ is a second triangle. Moreover, we have $(z_i, z_{i+1}) \in A(D)$ for all i, for otherwise $\{z_{i-1}, z_i, z_{i+1}\}$ is a triangle for some i. This implies that $z_{k-1}w \in A(D)$, a contradiction. Thus, some edge e_1 on P_1 is not oriented in D. If $C_2 - \{e_1\}$ is disconnected, let $S = \{e_1\}$. If $C_2 - \{e_1\}$ is not disconnected, then, by the same reasoning, there is a path P_2 in $C_2 - \{e_1\}$ connecting z and w such that some edge e_2 on P_2 is not oriented in D. Continuing in this way, we

will find a set $S = \{e_1, e_2, \cdots, e_t\}$, such that $e_i, i = 1, 2, \cdots, t$, is not oriented in D and $C_2 - S$ is disconnected. Note that since $e_i \in E(C_2)$, it is not a triangle edge. It follows that $e_i, i = 1, 2, \cdots, t$, must be taken care of by some added vertex a_i. Let $G_1 = G - S$. Then $\overline{E}_1 = S$. Let $I_{\overline{m}_1}$ be the set of $\overline{m}_1$ vertices in I_m that only take care of edges in $\overline{E}_1$. Then $\overline{m}_1 \geq t$. By Lemma 4, $p(G) \geq p(G_1) + \overline{m}_1 \geq p(G_1) + t$. Notice that G_1 is separable by z into two subgraphs H_1 and H_2, where $H_1 = C_1 \cup \{xz, yz\}$ is a connected graph with exactly one triangle x, y, z and $H_2 = C_2 - S$ is a graph with no triangles and at least two components. Thus, by Theorem 7, $p(H_1) = e(H_1) - n(H_1) - 1$, and by Lemma 5, $p(H_2) \geq e(H_2) - n(H_2) + 2$. It follows from Lemma 8 that $p(G_1) \geq p(H_1) + p(H_2) \geq (e(H_1) + e(H_2)) - (n(H_1) + n(H_2)) + 1 = (e - t) - (n + 1) + 1 = e - n - t$. Hence, $p(G) \geq e - n$, a contradiction. □

Lemma 24 *Let graph $G = (V, E)$ be separable by vertex w into two graphs G_1 and G_2. If in any optimal phylogeny digraph for G_1, we must have some edge of G_1 oriented toward w, and in any optimal phylogeny digraph for G_2, we must have some edge of G_2 oriented toward w, then $p(G) \geq p(G_1) + p(G_2) + 1$.*

Proof. Let $D = (V \cup I_m, A)$ be an optimal phylogeny digraph for G. By Lemma 8 we know that $D|_{G_1}$, $D|_{G_2}$ are phylogeny digraphs for G_1, G_2 respectively. Note that if $(u_1, w) \in A, (u_2, w) \in A$ for some $u_1 \in V_1$ and $u_2 \in V_2$, then $u_1u_2 \in E$, contradicting the fact that there is no edges between $V_1 - \{w\}$ and $V_2 - \{w\}$. So we can assume that no edges in G_2 is oriented toward w in $D|_{G_2}$, which implies that $D|_{G_2}$ is not an optimal phylogeny digraph for G_2 and so $p(G) = m = m_1 + m_2 \geq p(G_1) + p(G_2) + 1$. □

Lemma 25 *Let $G = (V, E)$ be a connected graph with exactly two triangles that are edge-disjoint. Assume the canonical notation for the triangles and assume that G^- has three components C_1, C_2 and C_3 with $z, w \in C_3$.*

(1) *If $x, y \in C_1$ and $u, v \in C_2$, or*

(2) *If $x, u \in C_1$ and $y, v \in C_2$,*

Then $p(G) = e - n - 2$.

Proof. To show that $p(G) \leq e - n - 2$, let $E' = E - \{uv, uw\}$ and let $G' = (V, E')$. Then under (1) or (2), we always have that G' is a connected graph with exactly one triangle, and the graph G'^- has two components. Using Theorem 7, $p(G') = (e - 2) - n - 1 = e - n - 3$. Let $D' = (V \cup I_{e-n-3}, A')$ be an optimal phylogeny digraph for G', and let D be the digraph obtained from D' by adding an isolated vertex a to V' and three arcs $(u, a), (v, a), (w, a)$ to A'. Then D is acyclic and is a phylogeny digraph for G with $|V(D)| - |V(G)| = |V(D')| + 1 - |V(G')| = p(G') + 1 = e - n - 2$. So $p(G) \leq e - n - 2$. We will show now that $p(G) \geq e - n - 2$.

Assume (1). It is easy to see that in this case G is separable by w into two subgraphs G_1 and G_2, where G_1 contains exactly one triangle x, y, z and G_2 contains exactly one triangle u, v, w. Now, G_1^-, the graph obtained from G_1 by deleting its three triangle edges, has two

components and G_2^- also has two components. By Theorem 7, $p(G_1) = e_1 - n_1 - 1$, $p(G_2) = e_2 - n_2 - 1$. By (1) of Lemma 23, for any optimal phylogeny digraph D_1 for G_1, $(x,z),(y,z) \in A(D_1)$. By (2) of Lemma 23, $(z_1,w) \in A(D_1)$ for some vertex z_1 of G_1. Similarly, for for any optimal phylogeny digraph D_2 for G_2, we must have that $(u,w),(v,w) \in A(D_2)$. Therefore, by Lemma 24, $p(G) \geq p(G_1)+p(G_2)+1 = (e_1-n_1-1)+(e_2-n_2-1)+1 = e-(n+1)-1 = e-n-2$.

Assume (2). The proof will be similar to that of Theorem 14. Let $D = (V \cup I_{p(G)}, A)$ be an optimal phylogeny digraph for G. Notice first that if there are two added ancestors a, b for the two triangles, then let $G_1 = G^-$. G_1 is a graph with three components and with no triangles, so by Lemma 5, $p(G_1) = (e-6) - n + 3 = e - n - 3$. Note that $\overline{E}_1$ is the six triangle edges, so that $a, b \in I_{\overline{m}_1}$. Since there is no mixed triangles, by Lemma 4, $p(G) \geq p(G_1) + \overline{m}_1 \geq e - n - 3 + 2 = e - n - 1$, contradicting $p(G) \leq e - n - 2$.

So there is at most one added ancestor for edges in the two triangles. Suppose there is one added ancestor a for edges in the two triangles. Without loss of generality, let us say a takes care of some edge in triangle u, v, w. Let $E_1 = E - \{uv, uw, vw\}$. Then $G_1 = (V, E_1)$ has one triangle with $G_1^- = G^-$ having three components. By Theorem 7, $p(G_1) = (e-3) - n = e - n - 3$. Since $\overline{E}_1$ is the three triangle edges for the triangle u, v, w, we have $a \in I_{\overline{m}_1}$. Since there is no mixed triangles, by Lemma 4, $p(G) \geq p(G_1) + \overline{m}_1 \geq e - n - 3 + 1 = e - n - 2$.

If there is no added ancestor for edges in the two triangles, then each of the triangles must have one vertex serving as a common ancestor for the triangle it is in.

- Suppose that at least one of z and w is the common ancestor for their corresponding triangle. Then we can assume without loss of generality that z is the common ancestor for triangle x, y, z, and either w or u is the common ancestor for triangle u, v, w. Let $E_1 = E - \{xy, zy, uv, wv\}$. Then $G_1 = (V, E_1)$ has no triangles and it has two components. By Lemma 5, $p(G_1) = (e-4) - n + 2 = e - n - 2$. Thus, by Lemma 13, $p(G) \geq p(G_1) + \overline{m}_1 \geq p(G_1) = e - n - 2$.

- Assume now that z is not the common ancestor for triangle x, y, z and w is not the common ancestor for triangle u, v, w. Then both xy and uv must be oriented in D. Let $E_1 = E - \{xz, yz, uw, vw\}$. Then $G_1 = (V, E_1)$ has no triangles and it has two components. By Lemma 5, $p(G_1) = (e-4) - n + 2 = e - n - 2$. Thus, by Lemma 13, $p(G) \geq p(G_1) + \overline{m}_1 \geq p(G_1) = e - n - 2$.

□

Lemma 26 *Let $G = (V, E)$ be a connected graph with exactly two triangles that are edge-disjoint. Assume the canonical notation for the triangles and assume that G^- has three components C_1, C_2, C_3. If $x, y, z, w \in C_3, u \in C_1$ and $v \in C_2$, then $p(G) = e - n - 2$.*

Proof. It is easy to see that in this case G is separable by w into two subgraphs G_1 and G_2, where G_1 contains exactly one triangle x, y, z and G_2 contains exactly one triangle u, v, w. Now, since G_1^-, the graph obtained from G_1 by deleting its three triangle edges, has one component and G_2^- has three components, by Theorem 7, $p(G_1) = e_1 - n_1, p(G_2) = e_2 - n_2 - 1$.

By Lemma 19, G_2 has an optimal phylogeny digraph that has no incoming arcs toward w. Using Lemma 9, $p(G) = p(G_1)+p(G_2) = (e_1-n_1)+(e_2-n_2-1) = e-(n+1)-1 = e-n-2$. □

Theorem 27 *Let $G = (V, E)$ be a connected graph with exactly two triangles that are edge-disjoint. Then*

$$p(G) = \begin{cases} e-n-1 & \text{if } G^- \text{ has five components} \\ e-n-2 & \text{if } G^- \text{ has four components,} \\ e-n-2 & \text{if } G^- \text{ has three components} \\ & \text{with each component containing exactly two triangle vertices,} \\ & \text{or with one component containing a triangle of } G, \\ e-n-3 & \text{otherwise} \end{cases}$$

Proof. For the cases when G^- has five, four, one or two components, the conclusion follows directly from Theorem 18, Theorem 22 and Theorem 20.

Assume now that G^- has three components C_1, C_2 and C_3 and use the canonical notation for the triangles. If C_i contains exactly two triangle vertices for each $i = 1, 2, 3$, then we can assume without loss of generality that $z, w \in C_3$. Thus we can assume that either $x, y \in C_1$, $u, v \in C_2$ or $x, u \in C_1$, $y, v \in C_2$. By Lemma 25, $p(G) = e-n-2$. If one component of G^-, say C_3, contains a triangle, say x, y, z, then $w \in C_3$. We can assume without loss of generality that $u \in C_1$ and $v \in C_2$. Thus $p(G) = e-n-2$ by Lemma 26.

Suppose that not every one of these three components contains exactly two triangle vertices, and that G^- has no components that contains a triangle of G. Then, we can assume without loss of generality that C_1 contains only one triangle vertex, say u. Then $|\{x, y, z\} \cap C_i| \neq \emptyset, i = 2, 3$. Consider the graph $G' = (V, E')$, where $E' = E - \{uv, wv\}$. Notice that $v \in C_2 \cup C_3$ is connected with other vertices in $C_2 \cup C_3$ by edges of triangle x, y, z. So G' is a connected graph with exactly one triangle whose vertex set is x, y, z. Notice also that G'^-, the graph obtained from G' by deleting its three triangle edges, has at most two components since in G'^-, via the edge uw, component C_1 is either connected with component C_2 when $w \in C_2$, or connected with component C_3 when $w \in C_3$. By Theorem 7, $p(G') = (e-2)-n-1 = e-n-3$. Let $D' = (V', A')$ be an optimal phylogeny digraph for G'. Then by Lemma 21, we can assume without loss of generality that $wu \in A'$ and u has no other incoming arcs in D'. Let D be a digraph obtained from D' by taking the same vertex set as D' and adding the arc (v, u) to A'. It is easy to see that D is still acyclic, and that D is a phylogeny digraph for G. Notice that $|V(D)|-|V(G)| = |V(D'|-|V(G')| = p(G') = e-n-3$, so $p(G) \leq e-n-3$. It follows from Theorem 20 that $p(G) = e-n-3$. □

Theorem 27 should be compared to the analogous result of Kim and Roberts [4], which says that if G is a connected graph with exactly two triangles that are edge-disjoint, then $k(G) = e-n, e-n-1$, or $e-n-2$.

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] R. D. Dutton and R. C. Brigham, A characterization of competition graphs, *Discrete Appl. Math.*, **6** (1983), 315–317.

[3] F. Harary, R.Z. Norman and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed graphs*, John Wiley, New York, 1965.

[4] S. R. Kim and F. S. Roberts, Competition numbers of graphs with a small number of triangles, *Discrete Appl. Math.*, **78** (1997), 153–162.

[5] 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 Biological and Social Sciences*, IMA Volumes in Mathematics and its Applications, Vol. **17**, Springer-Verlag, New York, 1989, 221–243.

[6] A. Raychaudhuri and F. S. Roberts, Generalized competition graphs and their applications, in P. Brucker and R. Pauly (eds.), *Methods of Operations Research*, Anton Hain, Konigstein, West Germany, 1985, 295–311.

[7] 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-Verlag, New York, 1978, 477-490.

[8] F. S. Roberts, *Graph Theory and its Applications to Problems of Society*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1978.

[9] F. S. Roberts, *Applied Combinatorics*, Prentice-Hall, Englewood Cliffs, NJ, 1984.

[10] F. S. Roberts and L. Sheng, Phylogeny numbers, *Discrete Applied Mathematics*, **87** (1998), 213–228.