



DIMACS EDUCATIONAL MODULE SERIES

MODULE 08-3 Centrality and Anticentrality in Trees Date prepared: May 2008

Atif Abueida University of Dayton Dayton, OH 45469 email: Atif.Abueida@notes.udayton.edu

Michael Ackerman Bellarmine University Louisville, KY 40205 email: mackerman@bellarmine.edu

Sul-Young Choi LeMoyne College Syracuse, NY 13214 email: SulYoung@lemoyne.edu

DIMACS Center, CoRE Bldg., Rutgers University, 96 Frelinghuysen Road, Piscataway, NJ 08854-8018 TEL: 732-445-5928 • FAX: 732-445-5932 • EMAIL: center@dimacs.rutgers.edu Web: http://dimacs.rutgers.edu/

Founded as a National Science Foundation Science and Technology Center and a Joint Project of Rutgers University, Princeton University, AT&T Labs - Research, Bell Labs, NEC Laboratories America and Telcordia Technologies with affiliated members Avaya Labs, Georgia Institute of Technology, HP Labs, IBM Research, Microsoft Research, Rensselaer Polytechnic Institute, and Stevens Institute of Technology.

Module Description Information

• Title:

Centrality and Anticentrality in Trees

• Authors:

Atif Abueida, University of Dayton Michael Ackerman, Bellarmine University Sul-Young Choi, LeMoyne College

• Abstract:

The idea of the "middle" of graphs has been studied extensively in the literature. In this module, we discuss different measures of the middle of trees. We also discuss the notion of the "antimiddle" of a tree. As part of our discussion of the "antimiddle" of a tree, we introduce a new measure of the "anti-middle" of a tree called the set of antiplurality vertices.

• Informal Description:

The study of the middle of graphs has numerous real-life applications. The focus of this module deals with the application of locating a facility such as a fire station or distribution center. It also examines where to locate hazardous facilities, such as a nuclear waste site, where the facility needs to be as far away as possible.

• Target Audience:

College students with discrete mathematics background.

• Prerequisites:

The module assumes knowledge of basic graph theory.

• Mathematical Field:

Graph Theory

• Applications Areas:

Facility location problems

• Mathematics Subject Classification:

Primary: 05C05; Secondary: 05C90

• Contact Information:

Michael Ackerman Department of Mathematics Bellarmine University Louisville, KY 40205 email: mackerman@bellarmine.edu

• Other DIMACS modules related to this module:

Module 03–2: Facility Location Problems

1 Introduction

One of the more interesting applications of graph theory involves facility location problems in which the following question is considered: where is the best place to locate a fire house, an ambulance station, a distribution center, a garbage dump, or a nuclear waste facility? In the first few situations, we wish to place the facility so that it is "close" to all possible destinations. In the latter two cases, we want the facility to be as far away as possible. In this module, we examine different ways in which these questions can be answered. For simplicity, we only use trees as our graph models; however, the ideas that we cover may also be applied to a more general graph. Section 2 deals with the question of locating a fire house, ambulance station, or distribution center, and we examine ways to find the "middle" of a tree. In Section 3 we consider the "antimiddle" of a tree to locate things such as a garbage dump or nuclear waste site.

This module is designed to have the students work the exercises and answer the questions posed as they proceed through the module. These can be done individually or discussed collectively as a class. It is important that the students give their best effort in completing the questions and exercises before continuing with the discussion in the module.

2 The "middle" of a tree

ZomZom University is building a new office for on-campus security. The administration at ZZU is concerned with the safety of any person in any building on campus. Thus, they decide to place the new security office so that it has the smallest response time to a complaint at any building. Where should the security office be placed?

We model the campus map using a graph where the vertices of the graph represent the buildings on campus, and the edges of the graph represent the walkways between buildings. For simplicity sake, we use a tree to depict the ZZU campus, which appears in Figure 1.

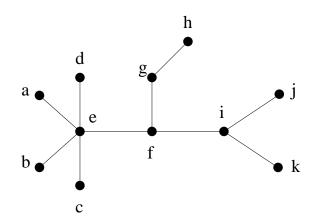


Figure 1: A tree model of ZZU's campus

Before we begin our discussion, we review some graph theoretic terms. The number of vertices in a graph G is called the **order** of G. A u-v **path** in a graph G is a finite, alternating sequence of vertices and edges of G, beginning at the vertex u and ending at the vertex v, in which no vertex is repeated. We typically denote a path using only the vertices in the sequence. For example, for the path a, (a, e), e, (e, f), f, (f, g), g in Figure 1, we simply write a e f g. A path whose beginning and ending vertices are the same is called a **cycle**. A graph which contains no cycles is a **tree**. If there is a path between every pair of distinct vertices in a graph G, we say that G is a **connected graph**. Otherwise, G is **disconnected**. A **component** of G is a subgraph of G that is connected and is not properly contained in any connected subgraph of G.

The **degree** of a vertex u in G, denoted deg(u), is the number of vertices to which u is adjacent in G. If deg(u) = 1, then u is called an **endvertex**. Finally, let u and v be vertices of G. The **distance between** u **and** v, denoted d(u, v), is the length of a shortest u-v path.

For example, referring to the connected graph of order 6 in Figure 2, e d c b is a path while e d c d a is not a path. The vertices e and f are endvertices, and a b c d a is a cycle. The degree of b is 3, and the distance between a and f is 2.

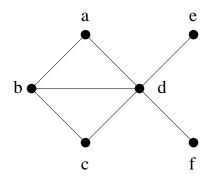


Figure 2: An example graph

Those wanting a more detailed discussion of graph theoretic terms are encouraged to consult Chartrand and Lesniak [3].

Returning to the tree that models ZZU at which vertex do you think the security office should be placed? Why?

Let us formally define the terms used in this process.

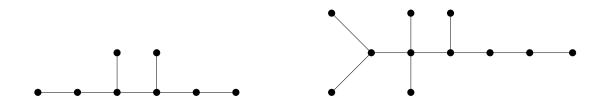
Let x be a vertex in the graph G. The eccentricity of x, denoted by e(x), is the distance to a vertex farthest from x. More formally, $e(x) = \max_{y \in V(T)} \{d(x, y)\}$. The radius of G, denoted by rad(G), is the minimum eccentricity of the vertices of G while the diameter of G, denoted by diam(G), is the maximum eccentricity of the vertices of G. Note that the length of the longest path in a tree is the diameter of the tree.

Go back and label each vertex in Figure 1 with its eccentricity.

After labelling the eccentricities of the vertices, did you get f as the vetex with smallest eccentricity? We refer to the set containing the vertex f as the center of the tree.

Definition 2.1. The set of vertices of a graph G with minimum eccentricity is called the **center** of G.

Exercise 2.2. Find the center of each tree given below.



Now, we consider a couple properties about the center of a tree. Additional properties of the center of a tree are left as exercises at the end of the section.

Proposition 2.3. The center of a tree T is contained on every longest path.

Notice that in the examples above, the center was either a single vertex or two adjacent vertices. This result is true for any tree. The proof of which is left as an exercise.

Theorem 2.4. The center of a tree T consists of one vertex or two adjacent vertices.

Now, we turn our attention back to the campus in our original example. A take-out Chinese restaurant receives permission to open a restaurant to exclusively serve the entire campus. Since the patrons on campus are very busy, the bulk of the restaurant's business will be from delivery service. Where should the restaurant be located? In this situation many trips will be made from the restaurant to the buildings on campus. Therefore, we want to minimize the sum of the distances from the restaurant to these buildings.

Definition 2.5. The status s(x) of a vertex x in a graph G is the sum of all distances from each $\left(\begin{array}{c} & & \\$

vertex of G to x
$$\left(1.e., s(x) = \sum_{y \in V(G)} a(x, y)\right)$$
.

Label each vertex of our campus tree in Figure 3 with its status.

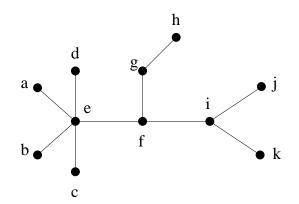


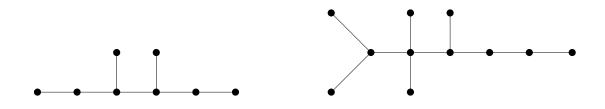
Figure 3: A tree model of ZZU's campus

We are looking for the vertices with smallest status.

Definition 2.6. The set of vertices of the graph G with minimum status is called the **median of** G.

So which set of vertices is the median of the tree in Figure 3?

Exercise 2.7. Find the median for the trees given below.



Notice that the center and the median of the right-most tree in the above exercise are disjoint. In fact, the center and median of a tree can be arbitrarily far apart.

As you probably discovered by now, calculating the status of each vertex of a tree T to find the median of T becomes quite complicated as the vertex set of T becomes larger. We now examine a procedure to find the median of a tree without having to calculate the status of a vertex. First, we need some preliminaries.

Definition 2.8. Let x and y be vertices in a tree T. Define V_{xy} to be the set of vertices of T that are closer to x than to y (i.e., $V_{xy} = \{z \in V(T) | d(x, z) < d(y, z)\}$).

For example, in the tree in Figure 3, $V_{ch} = \{c, b, a, d, e\}$ and $V_{hc} = \{h, g\}$. We have the following lemma.

Lemma 2.9. If xy is an edge in a tree T, then

$$s(x) = s(y) + |V_{yx}| - |V_{xy}|.$$

Exercise 2.10. Verify the status of vertex i in Figure 3 using the previous lemma.

Now, we discuss a procedure to find the median of a tree called the Majority Strategy, which was developed by H.M. Mulder in [7].

The Majority Strategy

- 1. Start at an initial vertex x_1 in the tree T.
- 2. For each $i \ge 1$, move from x_i to x_{i+1} where x_{i+1} is a vertex satisfying:
 - (i) x_i and x_{i+1} are adjacent,
 - (ii) $|V_{x_{i+1},x_i}| \ge |V_{x_i,x_{i+1}}|$, and

- (iii) $x_{i+1} \neq x_j$ for any j < i.
- 3. Stop when there are no more vertices into which to move.

Let x_p be the last vertex when the above process is applied. Then either

- (i) for all vertices v with $d(x_p, v) = 1$, $|V_{v,x_p}| < |V_{x_p,v}|$, or
- (ii) if v is a vertex $d(x_p, v) = 1$ and $|V_{v,x_p}| \ge |V_{x_p,v}|$, then $v = x_j$ for some j < p. In this case, $v_j = x_{p-1}$ since T does not contain a cycle.

In the first case above, the median is $\{x_p\}$ while $\{x_{p-1}, x_p\}$ is the median in the second case. From the examples that we have considered, it appears that the median of T is a single vertex or two adjacent vertices. Indeed, this is true in general.

Theorem 2.11. The median of a tree T consists of one single vertex or two adjacent vertices.

Proof. Use the Majority strategy discussed above. \blacksquare

We continue with another way to define the "middle" of a tree T in which we examine the size of the components of T after deleting a vertex.

Definition 2.12. Let T be a tree and let x be a vertex of T. The **branch weight of** x, denoted bw(x), is the number of vertices in a largest component of $T - \{x\}$.

Label each vertex in Figure 4 with its branch weight.

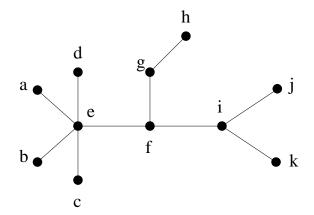


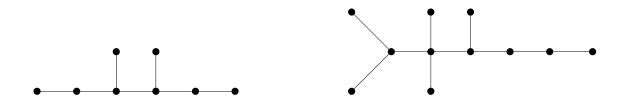
Figure 4: A tree model of ZZU's campus

We are looking for the set of vertices of smallest branch weight, which is called the centroid of T.

Definition 2.13. The set of all vertices of a tree T that have minimum branch weight is called the centroid of T.

In the example above, the centroid (T) is the set containing the vertex f.

Exercise 2.14. Find the centroid of the following trees.



How does the centroid for the trees in the previous examples compare to the median for each of those trees?

This observation leads us into our next theorem.

Theorem 2.15. In a tree T, the median of T equals the centroid of T.

Before we begin the proof of the theorem, we introduce two useful lemmas.

Lemma 2.16. Let xy be an edge in a tree. Then

- (i) if bw(x) < bw(y), then s(x) < s(y).
- (ii) if bw(x) = bw(y), then s(x) = s(y).

Proof. Let $\{B_{x1}, B_{x2}, ..., B_{xp}\}$ be the set of branches of x and $\{B_{y1}, B_{y2}, ..., B_{yq}\}$ the set of branches of y. Without loss of generality, we may suppose that B_{x1} contains y, and B_{y1} contains x. Then $V(B_{x1}) = V_{yx}, V(B_{y1}) = V_{xy}, B_{yi}$ is a subgraph of B_{x1} for all i > 1, and B_{xj} is a subgraph of B_{y1} for all j > 1.

Suppose bw(x) < bw(y), but $s(x) \ge s(y)$. From Lemma 2.9 and our assumption, $|V_{yx}| - |V_{xy}| = s(x) - s(y) \ge 0$. From the paragraph above, we have that for all j > 1, $|V(B_{xj})| < |V_{xy}| \le |V_{yx}| = |V(B_{x1})|$ and $bw(x) = |V(B_{x1})| = |V_{yx}|$. Since $bw(y) = max\{|V(B_{yi})| : i = 1, 2, ..., q\} \le max\{|V_{xy}|, |V_{yx}| - 1\} \le |V_{yx}|$, we have that $bw(y) \le bw(x)$, a contradiction. Therefore, bw(x) < bw(y) implies s(x) < s(y).

Suppose that bw(x) = bw(y), but $s(x) \neq s(y)$. Without loss of generality we may assume s(x) > s(y). Using a similar argument we can deduce that $bw(x) = |V_{yx}|$, $bw(y) < |V_{yx}|$, and bw(x) < bw(y). Therefore, bw(y) = bw(x) implies s(x) = s(y).

From the above lemma, we conclude that for two adjacent vertices x and y in a tree,

- (i) bw(x) < bw(y) if and only if s(x) < s(y), and
- (ii) bw(x) = bw(y) if and only if s(x) = s(y).

The converse of Lemma 2.12 is left as an exercise.

Lemma 2.17. Let x, y, and z be three vertices in a tree T such that d(x, z) = d(x, y) + d(y, z) = 2. Then:

- (i) $V_{xy} \subset V_{yz}$ and $V_{zy} \subset V_{yx}$
- (ii) If $s(x) \leq s(y)$, then s(y) < s(z).
- (iii) Also, if $bw(x) \leq bw(y)$, then bw(y) < bw(z).

Proof.

- (i) Exercise.
- (ii) Since xy is an edge in T, $s(x) s(y) = |V_{yx}| |V_{xy}|$ from Lemma 2.9, and thus, $|V_{yx}| \le |V_{xy}|$. From (i), $|V_{zy}| < |V_{yx}| \le |V_{yy}|$. So $s(y) - s(z) = |V_{zy}| - |V_{yz}| < 0$.
- (iii) From (ii) and previous Lemma.

As a result of the previous lemma, if $P = u_1 u_2 \ldots u_q$ is a path in a tree T and $s(u_1) \le s(u_2)$ (or $bw(u_1) \le bw(u_2)$), then $s(u_i) < s(u_j)$ (or $bw(u_i) < bw(u_j)$) for all i and j with i < j and j > 2.

Now, we are ready to prove the theorem.

Proof. Suppose that the median of T consists of two adjacent vertices x and y. Then s(x) = s(y) and bw(x) = bw(y). For a vertex u not in the median of T, since x and y are adjacent in a tree, there exists a path P containing three vertices x, y, and u. Without loss of generality we may assume that P is the x - u path. From the previous Lemma, s(x) < s(u) and bw(x) < bw(u). This imples that $Centroid(T) = \{x, y\}$.

The case that the median of T consists of only one vertex can be proven using a similar argument.

A consequence of the previous theorem is that the centroid of any tree consists of a single vertex or two adjacent vertices.

Next, we investigate our final way to determine the "middle" of a tree by pruning leaves from the tree in a particular order.

Definition 2.18. A processing sequence for a tree T of order n is a permutation $(x_1, x_2, ..., x_n)$ so that x_1 is an endvertex of T and for $2 \le j \le n$, x_j is an endvertex of $T - \{x_1, x_2, ..., x_{j-1}\}$.

Exercise 2.19. For the tree T in Figure 5, verify the following sequences are processing sequences for T:

$$P_{1} = (a, j, d, h, g, k, i, f, b, c, e)$$

$$P_{2} = (h, g, j, k, i, f, c, b, d, a, e)$$

$$P_{3} = (a, b, c, d, e, h, j, k, i, f, g)$$

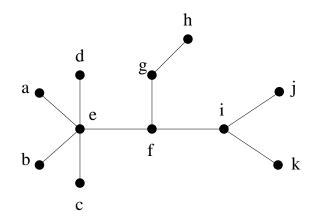


Figure 5: A tree model of ZZU's campus

Can you come up with a processing sequence in which the vertex e appears sooner than the fifth entry?

In fact, the fifth entry is the earliest that the vertex e can appear in a processing sequence.

Definition 2.20. Let x be a vertex in tree T. The processing number of x, denoted by proc(x), is the earliest entry at which x occurs in a processing sequence (i.e., $proc(x) = \min\{i|x = x_i \text{ in some processing sequence of } T\}$).

Definition 2.21. The processing center of T is the set of vertices of T with maximum processing number.

Find the processing number of each vertex and the processing center for the tree in Figure 6.

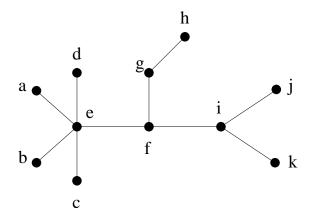


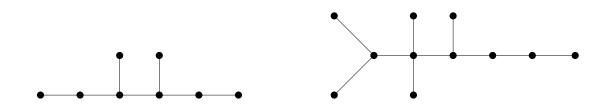
Figure 6: A tree model of ZZU's campus

How do the branch weights and processing numbers of each vertex in the tree above compare? This leads us to our next proposition. **Proposition 2.22.** For a vertex v in a tree T, proc(v) = |V(T)| - bw(v).

Proof. Suppose that (x_1, x_2, \ldots, x_n) is a processing sequence of a tree T. From the definition of processing sequence, the subgraph induced by $T - \{x_1, x_2, \ldots, x_{j-1}\}$ is connected. So the subgraph induced by $\{x_{j+1}, x_{j+2}, \ldots, x_n\}$ is a branch of x_j in T.

For a vertex v in T, let (x_1, x_2, \ldots, x_n) be a processing sequence corresponding to proc(v). Then $v = x_j$ for some j, and the subgraph induced by $\{x_{j+1}, x_{j+2}, \ldots, x_n\}$ is the largest branch of v in T. Thus, proc(v) = j = n - (n - j) = |V(T)| - bw(v).

Exercise 2.23. Find the processing center for the following trees.



How does the processing center found for the trees in the previous examples compare to the median(centroid)?

This leads us to our next theorem.

Theorem 2.24. In a tree T, the processing center of T is the same as the median of T.

Proof. From Proposition 2.22, we know that proc(u) = |V(T)| - bw(u) for each vertex u in T. By definition, the processing center of T is $\max_{u \in V(T)} \{proc(u)\} = \max_{u \in V(T)} \{|V(T)| - bw(u)\}$. Since |V(T)| is constant, the right-hand side of the equation above is maximized when bw(u) is minimized. Thus, the processing center of T is the same as the centroid of T. Therefore, the processing center of T is the same as the centroid of T.

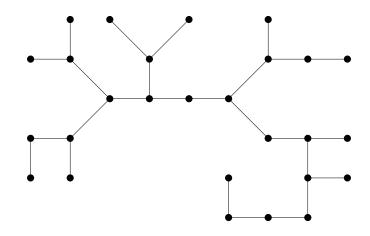
As a result of the previous theorem, the processing center of a tree T consists of a single vertex or two adjacent vertices.

Note that there are other ways to find the "middle" of a tree. If you would like to investigate some of these other ways, do not hesitate to visit your friendly professor.

Additional Exercises

- 1. Let G be a connected graph. Prove that $rad(G) \leq diam(G) \leq 2rad(G)$. For each inequality above, give an example of a graph of order greater than 2 where equality holds.
- 2. Let r and d be fixed positive integers where $r \leq d \leq 2r$. Construct a graph with radius r and diameter d.
- 3. Let x and y be vertices of the graph G. Prove that $|e(x) e(y)| \le d(x, y)$.

- 4. Show that the center of a tree T is contained on every longest path.
- 5. Find the center, median, centroid, and processing center for the given tree.



- 6. Prove that the center of a tree is a single vertex or two adjacent vertices.
- 7. Show that a tree T has just a single vertex as its center if and only if diam(T) = 2rad(G).
- 8. Construct a tree in which the center and median are arbitrarily far apart. Can this be done so that both contain two vertices?
- 9. Prove Lemma 2.9.
- 10. Let xy be an edge in a tree T. Prove that: (a) if s(x) < s(y), then bw(x) < bw(y)(b) if s(x) = s(y), then bw(x) = bw(y)
- 11. Prove Lemma 2.17 (i).

3 The "antimiddle" of trees

Returning the beautiful campus of ZomZom University, they now have quite an accumulation of garbage that is making the campus unattractive. So, the administration approved a huge dumpster directly behind one of the buildings on campus to collect all of the campus' garbage. Where should the dumpster be located?

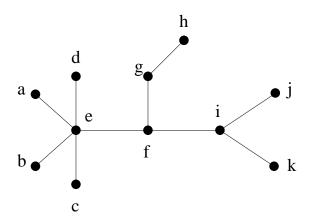


Figure 7: A tree model of ZZU's campus

In this application we want to put the dumpster "as far away" as possible. Initially, we are interested in those vertices having maximum eccentricity.

Definition 3.1. Let T be a tree. The **anticenter of** T is the set of vertices with maximum eccentricity.

From the definition of anticenter, we obtain the following proposition.

Proposition 3.2. The anticenter of a tree T is the set of end vertices on the longest paths of T.

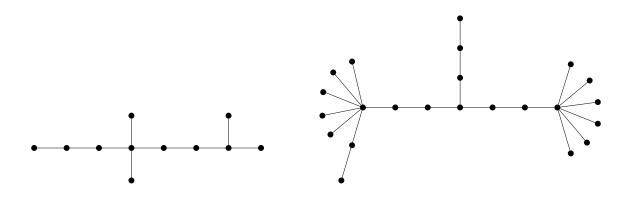
If we refer to the eccentricities of the vertices that we found in Section 2, the anticenter of T is $\{a, b, c, d, h, j, k\}$. However, this poses a problem. How do we decide among these 7 vertices where to put the huge dumpster?

Let us examine another notion of "anticentrality". In this situation we are looking for those vertices whose status is as large as possible.

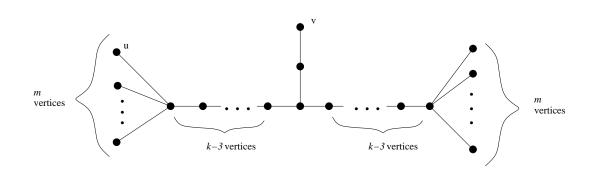
Definition 3.3. Let T be a tree. The antimedian of T is the the set of vertices with maximum status.

Consulting our previous work in which we found the status of each vertex in the tree in Figure 7, we see that the vertex h has the maximum status. Thus, the antimedian of T is $\{h\}$, and in this case, the antimedian makes the decision of where to place the dumpster much easier.

Exercise 3.4. Find the anticenter and antimedian of the trees below.



From the previous exercise, we notice that an antimedian vertex of a tree is not necessarily on a longest path. Thus, an antimedian vertex is not necessarily an anticenter vertex. A natural question arises: How do the antimedian and anticenter of a tree compare? In [4], Johns shows that they can be made arbitrarily far apart. We omit the proof of the theorem but illustrate the idea using the tree below where $k \ge 5$ and $m = \frac{k^2 - 5k + 10}{2}$.



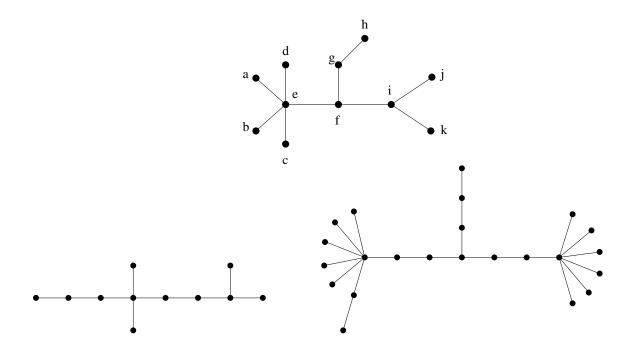
In the tree above, one should verify that e(v) = k and e(u) = 2k - 4. So the anticenter contains all 2k vertices at either end of the tree. However, $s(v) = k^3 - 4k^2 + 9k - 3$ and $s(u) = k^3 - 4k^2 + 8k - 3$. Thus, the antimedian contains only the vertex v.

We employ one other strategy to determine "antimiddle" vertices in a tree.

Definition 3.5. Let T be a tree. The vertex x is an **antiplurality vertex of** T if and only if $|V_{xy}| \leq |V_{yx}|$ for all y in V(T).

It is obvious from the definition that an antiplurality vertex of a tree is an end vertex.

Exercise 3.6. Find all antiplurality vertices for the trees below.



For the trees in the previous exercise, there does not appear to be any obvious relationship between the antimedian and the set of antiplurality vertices in a tree. In fact, an antimedian vertex and an antiplurality vertex can be arbitrarily far apart (See Exercise 3 in this section). However, by examining the examples above, it seems that an antiplurality vertex is an anticenter vertex. The next proposition confirms that notion.

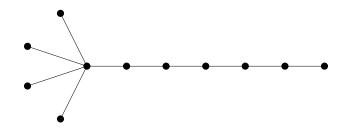
Proposition 3.7. In a tree T, an antiplurality vertex of T is an endvertex on a longest path in T

Proof. Suppose that u is an antiplurality vertex but not an endvertex on any longest path in the tree T. Let P be a longest path in T with endvertices x and y, and let w be the vertex on P such that $d(u, w) = \min\{d(u, v)|v \text{ is on the } u\text{-}x \text{ path}\}$. Since P is a longest path in T, d(u, w) < d(x, w), and d(u, w) < d(w, y). This means that $V_{xu} \subsetneq V_{uy}$ and $V_{yu} \subsetneq V_{ux}$. Since u is an antiplurality vertex, $|V_{ux}| \le |V_{xu}|$. So we have that $|V_{yu}| < |V_{ux}| \le |V_{xu}|$ which contradicts the fact that u is an antiplurality vertex. Thus, u is an endvertex on a longest path in T.

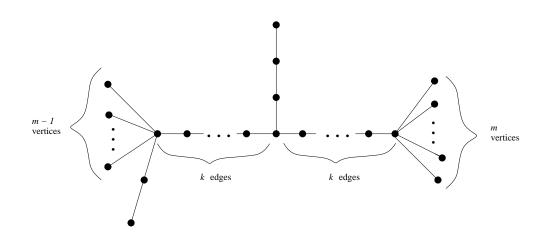
There are other ways to determine the "antimiddle" of a tree T. If you would like to investigate some of these other ways, do not hesitate to visit your friendly professor.

Additional Exercises

1. Find the anticenter, antimedian and all antiplurality vertices for the given tree.



- 2. Construct a tree of order at least 5 whose anticenter, antimedian, and antiplurality vertices are the same.
- 3. For the following tree, determine the anticenter, antimedian and antiplurality vertices for several values of k and m. Find the value of k and m for which the distance between an antimedian vertex and an antiplurality vertex (or an anticenter vertex) is at least 9.



References

- [1] F. Buckley, "Facility location problems", College Math. J. 18 (1987), 24-32.
- [2] F. Buckley and F. Harary, Distance in Graphs, Addison-Wesley Publ. Co, New York (1990).
- [3] G. Chartrand and L. Lesniak, *Graphs and Digraphs*, Chapman & Hall, London (1996).
- [4] G. Johns, "Margins in graphs and trees", Congr. Numer. 134 (1998), 99-105.
- [5] M. Labbe, "Location of an obnoxious facility on a network: a voting approach", Networks 20 (1990), 197-207.
- [6] E. Minieka, "Anticenters and antimedians of a network", Networks 13 (1983), 359-364.
- [7] H.M.Mulder, "The majority strategy on graphs", Discrete Applied Math. 80 (1997), 97-105.
- [8] B. Zelinka, "Medians and peripherians of trees", Arch. Math. (Brno) 4 (1968), 87-95.