Bohdan Zelinka

Distances between rooted trees

*Mathematica Bohemica*, Vol. 116 (1991), No. 1, 101–107

Persistent URL: http://dml.cz/dmlcz/126191

# DISTANCES BETWEEN ROOTED TREES

BOHDAN ZELINKA, Liberec
(Received May 12, 1989)

*Summary.* Two types of a distance between isomorphism classes of graphs are adapted for rooted trees.

*Keywords and phrases:* rooted tree, subtree distance, edge rotation distance.

*Classification:* 05C05.

Various distances between isomorphism classes of graphs were introduced and studied by various authors. In this paper we shall adapt the definitions of two of them for rooted trees and study their properties. We consider finite trees. The symbol $V(T)$ denotes the vertex set of the tree $T$.

Let $T_1, T_2$ be two trees with the same number $n$ of vertices. The subtree distance $\delta_T(T_1, T_2)$ is equal to $n$ minus the maximum number of vertices of a tree which is isomorphic simultaneously to a subtree of $T_1$ and to a subtree of $T_2$. This distance was introduced by the author of this paper in [2].

Let $u, v, w$ be three pairwise distinct vertices of an undirected graph $G$ such that $u$ is adjacent to $v$ and is not adjacent to $w$. Let $e$ be the edge joining $u$ and $v$. To perform a rotation of the edge $e$ to the position $uw$ means to delete $e$ from $G$ and add the edge $uw$ to $G$.

The edge rotation distance was introduced by G. Chartrand, F. Saba and H.-B. Zou in [1]. Let $G_1, G_2$ be two finite undirected graphs with the same number of vertices and the same number of edges. In [1] it was proved that $G_1$ can be transformed to a graph isomorphic to $G_2$ by a finite number of edge rotations. The minimum number of edge rotations necessary for it is called the edge rotation distance $\delta_R(G_1, G_2)$ of the graphs $G_1$ and $G_2$.

Both the above described distances are metrics. If we wanted to speak precisely, we should speak about the distance between isomorphism classes of graphs instead of the distance between graphs. Namely, if two graphs have the distance equal to zero, they need not be identical, but they are isomorphic. Nonetheless, for the sake of simplicity and brevity we shall speak about distances between graphs.

We shall adapt these definitions for rooted trees.

A rooted tree is an ordered pair $(T, r)$, where $T$ is a tree and $r$ is a vertex of $T$; the vertex $r$ is called the root of the rooted tree $(T, r)$. Two rooted trees $(T_1, r_1)$, $(T_2, r_2)$ are called isomorphic, if there exists an isomorphic mapping of $T_1$ onto $T_2$

which maps $r_1$ onto $r_2$. When we speak about the number of vertices of a rooted tree $(T, r)$, we mean the number of vertices of $T$. A rooted subtree of $(T, r)$ is a rooted tree $(T_0, r)$, where $T_0$ is a subtree of $T$ containing the vertex $r$.

Let $(T_1, r_1)$, $(T_2, r_2)$ be two rooted trees with the same number $n$ of vertices. The distance $\delta_T((T_1, r_1), (T_2, r_2))$ is equal to $n$ minus the maximum number of vertices of a rooted tree which is isomorphic simultaneously to a rooted subtree of $(T_1, r_1)$ and to a rooted subtree of $(T_2, r_2)$.

**Theorem 1.** *The distance $\delta_T$ is a metric on the set of all isomorphism classes of rooted trees with n vertices for any positive integer n.*

Proof. It is clear that $\delta_T((T_1, r_1), (T_2, r_2)) \geqq 0$ for any two rooted trees $(T_1, r_1)$ $(T_2, r_2)$ with the same number of vertices, and $\delta_T((T_1, r_1), (T_2, r_2)) = 0$ if and only if $(T_1, r_1) \cong (T_2, r_2)$. Also it is evident that $\delta_T((T_1, r_1), (T_2, r_2)) = \delta_T((T_2, r_2), (T_1, r_1))$. It remains to prove the triangle inequality. Let $(T_1, r_1)$, $(T_2, r_2)$, $(T_3, r_3)$ be three rooted trees with $n$ vertices each. There exists a rooted tree $(T_{12}, r_{12})$ with $n - \delta_T((T_1, r_1), (T_2, r_2))$ vertices which is simultaneously isomorphic to a rooted subtree of $(T_1, r_1)$ and to a rooted subtree of $(T_2, r_2)$, and a rooted tree $(T_{23}, r_{23})$ with $n - \delta_T((T_2, r_2), (T_3, r_3))$ vertices which is simultaneously isomorphic to a rooted subtree of $(T_2, r_2)$ and a rooted subtree of $(T_3, r_3)$. We may consider $(T_{12}, r_{12})$ and $(T_{23}, r_{23})$ directly as rooted subtrees of $(T_2, r_2)$ with $r_{12} = r_{23} = r_2$. Let $T_{13}$ be the intersection of the trees $T_{12}$, $T_{23}$; it contains $r_2$ and thus it has at least one vertex. The rooted tree $(T_{13}, r_2)$ is evidently isomorphic to a rooted subtree of $(T_1, r_1)$ and to a rooted subtree of $(T_3, r_3)$ and thus

$$(1) \qquad \delta_T((T_1, r_1), (T_3, r_3)) \leqq n - |V(T_{13})| .$$

Let $T'_{13}$ be the union of $T_{12}$ and $T_{23}$. Then

$$|V(T'_{13})| = |V(T_{12})| + |V(T_{23})| - |V(T_{13})| =$$
$$= 2n - \delta_T((T_1, r_1), (T_2, r_2)) - \delta_T((T_2, r_2), (T_3, r_3)) - |V(T_{13})| .$$

As $T'_{13}$ is a subtree of $T_2$, we have $|V(T'_{13})| \leqq n$. This yields

$$2n - \delta_T((T_1, r_1), (T_2, r_2)) - \delta_T((T_2, r_2), (T_3, r_3)) - |V(T_{13})| \leqq n ,$$

i.e.

$$n - |V(T_{13})| \leqq \delta_T((T_1, r_1), (T_2, r_2)) + \delta_T((T_2, r_2), (T_3, r_3)) . \qquad \square$$

Together with (1) this implies the triangle inequality for $\delta_T$.

Let again $(T_1, r_1)$, $(T_2, r_2)$ be two rooted trees with the same number $n$ of vertices. The edge rotation distance $\delta_R((T_1, r_1), (T_2, r_2))$ between these rooted trees is the minimum number of edge rotations necessary for transforming $(T_1, r_1)$ to a rooted tree isomorphic to $(T_2, r_2)$, where all graphs occurring at performing these rotations are trees and are considered as rooted trees with the same root $r_1$.

To show that the edge rotation distance between two rooted trees is well-defined, we shall prove a theorem.

**Theorem 2.** *Let $(T_1, r_1)$, $(T_2, r_2)$ be two rooted trees with the same number n of vertices. Then there exists a finite number of edge rotations which transform $(T_1, r_1)$ to a rooted tree isomorphic to $(T_2, r_2)$ and this number is less than or equal to $\delta_T((T_1, r_1), (T_2, r_2))$.*

Proof. We shall use induction with respect to the distance $\delta_T((T_1, r_1), (T_2, r_2))$. If $\delta_T((T_1, r_1), (T_2, r_2)) = 0$, then evidently the rooted trees $(T_1, r_1)$, $(T_2, r_2)$ are isomorphic and the assertion holds trivially. Let $k \geqq 1$ and suppose that the assertion is true for any two rooted trees whose subtree distance is at most $k - 1$. Let $\delta_T((T_1, r_1), (T_2, r_2)) = k$. Then there exists a rooted tree $(T_0, r_0)$ with $n - k$ vertices which is isomorphic simultaneously to a rooted subtree $(T_{01}, r_1)$ or $(T_1, r_1)$ and to a rooted subtree $(T_{02}, r_2)$ of $(T_2, r_2)$. As $n - k < n$, there exists a vertex $x_2$ of $T_2$ which does not belong to $T_{02}$ and is adjacent in $T_2$ to a vertex $y_2$ of $T_{02}$. Let $\alpha$ be an isomorphic mapping of $T_{02}$ onto $T_{01}$ such that $\alpha(r_2) = r_1$, let $y_1 = \alpha(y_2)$. Let $x_1$ be a vertex of $T_1$ which does not belong to $T_{01}$ and is adjacent to a vertex $z$ of $T_{01}$. We perform the rotation of the edge $x_1z$ to the position $x_1y_1$. Let $T'_{01}$ (or $T'_{02}$) be the tree obtained from $T_{01}$ (or $T_{02}$) by adding the edge $x_1y_1$ (or $x_2y_2$, respectively). Let $\alpha'$ be the mapping of $T'_{02}$ onto $T'_{01}$ defined in such a way that $\alpha'(x_2) = x_1$ and $\alpha'(v) = \alpha(v)$ for each vertex $v$ of $T_{02}$. Then $\alpha'$ is an isomorphic mapping of $T'_{02}$ onto $T'_{01}$. Let $T'_1$ be the tree obtained from $T_1$ by the rotation of $x_1z$ to the position $x_1y_1$. Each of the trees $T'_{01}$, $T'_{02}$ has $n - (k - 1)$ vertices. We have $\alpha'(r_2) = r_1$. Thus the rooted trees $(T'_{01}, r_1)$, $(T'_{02}, r_2)$ are isomorphic and therefore $\delta_T((T_1, r_1), (T_2, r_2)) \leqq k - 1$. By the induction hypothesis there exists a finite number of edge rotations which transform $(T'_1, r_1)$ into $(T_2, r_2)$ and this number is at most $k - 1$. The rooted tree $(T'_1, r_1)$ is obtained from $(T_1, r_1)$ by the rotation of the edge $x_1z$ to the position $x_1y_1$, and this proves the assertion. □

**Corollary 1.** *For any two rooted trees $(T_1, r_1)$, $(T_2, r_2)$ with the same number of vertices the inequality*

$$\delta_R((T_1, r_1), (T_2, r_2)) \leqq \delta_T((T_1, r_1), (T_2, r_2))$$

*holds.* □

Let $(T, r)$ be a rooted tree. By $\Delta(T, r)$ we denote the degree of the root $r$ in $T$, by $D(T, r)$ we denote the maximum distance of a vertex of $T$ from $r$.

**Theorem 3.** *Let $(T_1, r_1)$, $(T_2, r_2)$ be two rooted trees with the same number of vertices. Then*

$$\delta_R((T_1, r_1), (T_2, r_2)) \geqq |\Delta(T_1, t_1) - \Delta(T_2, r_2)|,$$

$$\delta_T((T_1, r_1), (T_2, r_2)) \geqq |\Delta(T_1, r_1) - \Delta(T_2, r_2)|.$$

Proof. Consider edge rotations which transform $(T_1, r_1)$ to a rooted tree isomorphic to $(T_2, r_2)$. At each of them the degree of $r$ can change at most by one and therefore the number of such rotations is at least $\Delta(T_1, r_1) - \Delta(T_2, r_2)$, which implies the inequality for $\delta_R$. The inequality for $\delta_T$ follows from this inequality and from Theorem 2. □

**Theorem 4.** *Let $(T_1, r_1)$, $(T_2, r_2)$ be two rooted trees with the same number of vertices. Then*

$$\delta_T((T_1, r_1), (T_2, r_2)) \geqq |D(T_1, r_1) - D(T_2, r_2)| .$$

Proof. Without loss of generality we may suppose that $D(T_1, r_1) \geqq D(T_2, r_2)$. Let $v$ be a vertex of $T_1$ whose distance from $r_1$ is equal to $D(T_1, r_1)$. Let $P$ be the path connecting $r_1$ and $v$ in $T_1$, let $w$ be the vertex of $P$ whose distance from $r_1$ is equal to $D(T_2, r_2)$. Let $P'$ be the path in $T_1$ connecting $v$ and $w$. For any rooted subtree of $(T_2, r_2)$ the value of $D$ is at most $D(T_2, r_2)$; therefore no rooted subtree of $(T_1, r_1)$ isomorphic to a rooted subtree of $(T_2, r_2)$ can contain the vertices of $P'$ except $w$. The number of these vertices is $D(T_1, r_1) - D(T_2, r_2)$, which implies the assertion. □

Now we shall consider pairs $(T, r_1)$, $(T, r_2)$ of rooted trees, where the tree is the same and the roots are different.

Let $T$ be a tree with $n$ vertices. The maximum distance $\delta_T((T, r_1), (T, r_2))$, where $r_1, r_2$ are two vertices of $T$, will be called the subtree elongation of $T$ and denoted by $\varepsilon_T(T)$. The maximum distance $\delta_R((T, r_1), (T, r_2))$, where $r_1, r_2$ are vertices of $T$, will be called the edge rotation elongation of $T$ and denoted by $\varepsilon_R(T)$.

**Theorem 5.** *Let $T$ be a tree with $n \geqq 3$ vertices. Then*

$$[n/2] - 1 \leqq \varepsilon_T(T) \leqq n - 2 .$$

*The lower bound is attained for a snake, the upper bound for a star.*

Proof. Consider the decompositions of $T$ into two edge-disjoint subtrees having exactly one vertex in common. Let $\{T_1, T_2\}$ be such a decomposition with the property that the absolute value of the difference $|V(T_1)| - |V(T_2)|$ is minimum. Let $u$ be the common vertex of $T_1$ and $T_2$. (In [3] it is proved that $u$ is a median of $T$, i.e. a vertex having the minimum sum of distances from all vertices of $T$.) We shall prove that no branch of $T$ at the vertex $u$ has more than $[n/2] + 1$ vertices. Without loss of generality suppose that $|V(T_1)| \geqq |V(T_2)|$. We evidently have $|V(T_1)| + |V(T_2)| = n + 1$. Therefore if $|V(T_1)| - |V(T_2)| \leqq 1$, then $|V(T_2)| \leqq |V(T_1)| \leqq [n/2] + 1$. As each branch of $T$ at the vertex $u$ is a subtree of $T_1$ or $T_2$, the assertion holds. Thus let $|V(T_1)| - |V(T_2)| \geqq 2$. Suppose that a branch $B$ of $T$ at $u$ with more than $[n/2] + 1$ vertices exists. Let $u'$ be its vertex adjacent to $u$. It is clear that $B$ is a subtree of $T_1$. Let $T_1'$ be the tree obtained from $B$ by deleting the vertex $u$ and the edge $uu'$.

Let $T_2'$ be the subtree of $T$ formed by all edges not belonging to $T_1'$ and by all vertices incident to them. The tree $T_1'$ is a proper subtree of $T_1$ and $T_2$ is a proper subtree of $T_2'$. Hence $|V(T_1')| \leqq |V(T_1)| - 1$, $|V(T_2')| \geqq |V(T_2)| + 1$. The pair $\{T_1', T_2'\}$ is a decomposition of $T$ into two edge-disjoint subtrees with the unique common vertex $u'$. As $|V(B)| \geqq [n/2] + 1$, we have $|V(T_1')| \geqq [n/2] + 1 \geqq [n/2]$, $|V(T_2')| = n + 1 - |V(T_1')| \leqq [n/2]$. Hence $0 \leqq |V(T_1')| - |V(T_2')| \leqq |V(T_1)| - (|V(T_2)| - 2$, which is a contradiction with the minimality of the absolute value of $|V(T_1)| - |V(T_2)|$. We have proved that any branch of $T$ at $u$ has at most $[n/2] + 1$ vertices. Let $v$ be a terminal vertex of $T$ and consider rooted trees $(T, u)$, $(T, v)$. Let $T_0$ be a subtree of $T$ containing $u$ and such that $(T_0, u)$ is isomorphic to a rooted subtree of $T$ with the root $v$. As the degree of $v$ is 1, the tree $T_0$ is a subtree of one branch of $T$ at $u$ and $|V(T_0)| \leqq [n/2] + 1$. Hence $\varepsilon_T(T) \geqq \delta_T((T, u), (T, v)) \geqq n - ([n/2] + 1) = [n/2] - 1$. On the other hand, a rooted tree with two vertices is a rooted subtree of every rooted tree with at least two vertices, which implies $\varepsilon_T(T) \leqq n - 2$.

Let $T$ be a snake with $n$ vertices. Then at each vertex of $T$ there exists a branch with at least $[n/2] + 1$ vertices and this branch is a snake. This implies $\varepsilon_T(T) = [n/2] - 1$.

Let $T$ be a star with $n$ vertices, let $c$ be its center, let $v$ be a vertex of $T$ different from $c$. Then $\delta_T((T, c), (T, v)) = n - 2$, because any branch of $T$ at $c$ has only two vertices. This implies $\varepsilon_T(T) = n - 2$. □

**Theorem 6.** *Let $k, n$ be two positive integers such that $[n/2] - 1 \leqq k \leqq n - 2$. Then there exists a tree $T$ with $n$ vertices such that $\varepsilon_T(T) = k$.*

Proof. Denote $p = n - k - 1$. Take a snake of length $2p$, i.e. with $2p + 1$ vertices. It has the unique center $c$. Add $n - 2p - 1$ new vertices and join each of them by an edge with $c$. The resulting tree is $T$. Let $v$ be a terminal vertex of the above mentioned snake and consider the rooted trees $(T, c)$, $(T, v)$. Evidently the rooted tree isomorphic to subtrees of both of them and having the maximum number of vertices is a snake with $p + 1$ vertices and thus $\delta_T((T, c), (T, v)) = n - p - 1 = k$. Any vertex of $T$ different from $c$ has the property that it is a terminal vertex of a snake of length at least $p + 2$ which is a subtree of $T$ and thus the distance of any two rooted trees $(T, r_1)$, $(T, r_2)$ is at most $k$ and $\varepsilon_T(T) = k$. □

**Theorem 7.** *Let $k, n$ be two positive integers such that $1 \leqq k \leqq n - 2$. Then there exists a tree $T$ with $n$ vertices such that $\varepsilon_R(T) = k$.*

Proof. Take a snake with $n - k$ vertices $u_1, \ldots, u_{n-k-1}, v$ and edges $u_i u_{i+1}$ for $i = 1, \ldots, n - k - 2$ and $u_{n-k-1}v$. Add $k$ new vertices $w_1, \ldots, w_k$ and join them by edges with $v$. The resulting tree is $T$. According to Theorem 3 we have $\delta_R((T, v), (T, x)) \geqq k$ for any vertex $x$ of $T$ of degree 1, because $v$ has degree $k + 1$. By $k$ rotations of edges $yv$ to the position $yu_1$, where $y$ is a vertex of degree 1 different from $u_1$, we transform $(T, v)$ to a rooted tree isomorphic to $(T, u_1)$. We can transform

$(T, v)$ to a rooted tree isomorphic to $(T, w_1)$ for $i = 1, \ldots, k$ by $k - 1$ rotations of edges $w_j v$ to the position $w_j w_i$, where $j \neq i$, and by the rotation of $u_1 u_2$ to the position $u_1 w_i$. Thus $\delta_R((T, v), (T, u_1) = \delta_R((T, v), (T, w_i)) = k$ for $i = 1, \ldots, k$. If $u_2$ exists, we can transform $(T, v)$ to $(T, u_2)$ by $k - 1$ rotations of $w_i v$ to the position $w_i u_2$ for $i = 1, \ldots, k - 1$. This and Theorem 3 imply $\delta_R((T, v), (T, u_2)) = k - 1$. If $u_i$ exists for $i \geqq 3$, then we can transform $(T, v)$ to a rooted tree isomorphic to $(T, u_i)$ by $k - 1$ rotations of $w_j v$ to the position $w_j u_i$ for $j = 1, \ldots, k - 1$ and by the rotation of $u_{i-2} u_{i-1}$ to the position $u_{i-2} w_k$; therefore $\delta_R((T, v), (T, u_i)) \leqq k$ for $i \geqq 3$. We can transform $(T, u_i)$ for each $i \geqq 2$ to a rooted tree isomorphic to $(T, u_1)$ by the rotation of $u_{i+1} u_i$ (or $v u_i$, if $i = n - k - 1$) to the position $u_{i+1} u_1$ (or $v u_1$). This implies also that for any $i, j$ from the numbers $2, \ldots, n - k - 1$ the rooted tree $(T, u_i)$ can be transformed to a rooted tree isomorphic to $(T, u_j)$ by two edge rotations; the intermediate rooted tree is $(T, u_1)$. The rooted tree $(T, u_1)$ can be transformed to a rooted tree isomorphic to $(T, w_i)$ for each $i \in \{1, \ldots, k\}$ by $k - 1$ edge rotations; they are the rotations of $w_j v$ to the position $w_j u_2$ for all $j \in \{1, \ldots, k\} - \{i\}$. Finally, any rooted tree $(T, u_j)$ for $j \geqq 2$ can be transformed to a rooted tree isomorphic to $(T, u_1)$ by one edge rotation, and thus to a rooted tree isomorphic to $(T, w_i)$ for any $i \in \{1, \ldots, k\}$ by $k$ edge rotations. We have proved that any two rooted trees resulting from $T$ have the distance at most $\max(k, 2)$. This proves our assertion in the case when $k \geqq 2$. In the case $k = 1$ the tree $T$ is a snake. We may denote its vertices more simply by $u_1, \ldots, u_n$ in such a way that its edges are $u_i u_{i+1}$ for $i = 1, \ldots, n - 1$. Let $u_i, u_j$ be two vertices of $T$ such that $i > j$. Then $(T, u_j)$ can be transformed to a rooted tree isomorphic to $(T, u_i)$ by the rotation of $u_{j-i} u_{j-i+1}$ to the position $u_{j-i} u_n$. Hence in this case $\varepsilon_T(T) = 1 = k$. □

In the end we formulate two corollaries.

**Corollary 2.** *Let $T$ be a finite tree, let $\Delta$ be the maximum degree of a vertex of $T$. Then*

$$\varepsilon_T(T) \geqq \varepsilon_R(T) \geqq \Delta - 1 .$$

Proof. The inequality $\varepsilon_T(T) \geqq \varepsilon_R(T)$ follows from Corollary 1, the inequality $\varepsilon_R(T) \geqq \Delta - 1$ from Theorem 3, because every finite tree contains a vertex of degree 1. □

**Corollary 3.** *Let $T$ be a tree with $n$ vertices, let $d$ be its diameter. Then*

$$\varepsilon_T(T) \geqq [d/2] .$$

Proof. Let $c$ be a center of $T$, let $v$ be a terminal vertex of a diametral path in $T$. We have $D(T, v) = d$, $D(T, c) = [d/2]$ (the radius of $T$). Hence the inequality follows from Theorem 4. □

*References*

[1] *G. Chartrand, F. Saba, H.-B. Zou:* Edge rotations and distance between graphs. Časopis pěst. mat. *110* (1985), 87—91.
[2] *B. Zelinka:* A distance between isomorphism classes of trees. Czechoslovak Math. J. *33* (1983), 126—130.
[3] *B. Zelinka:* Medians and peripherians of trees. Arch. Math. Brno *4* (1968), 87—95.

## Souhrn

## VZDÁLENOSTI MEZI ZAKOŘENĚNÝMI STROMY

BOHDAN ZELINKA

Dva typy vzdálenosti mezi třídami isomorfismu stromů jsou upraveny na vzdálenosti mezi třídami isomorfismu zakořeněných stromů. Zkoumají se jejich základní vlastnosti.

*Author's address:* Katedra matematiky VŠST, Sokolská 8, 460 01 Liberec.