ELSEVIER

Discrete Mathematics 161 (1996) 121–132

DISCRETE MATHEMATICS

Distances between graphs under edge operations

Wayne Goddard [*,1], Henda C. Swart

University of Natal, Durban 4000, South Africa

Received 9 September 1991; revised 18 October 1994

Abstract

We investigate three metrics on the isomorphism classes of graphs derived from elementary edge operations: the edge move, rotation and slide distances. We derive relations between the metrics, and bounds on the distance between arbitrary graphs and between arbitrary trees. We also consider the sensitivity of the metrics to various graph operations.

1. Introduction

One of the most fundamental concepts in graphs is that of isomorphism. Much time has been spent on how to decide whether two graphs are the same. But comparatively little work has been done on comparing how much different two graphs are. Of course, one may compare the values of various parameters for the two graphs. But what one really wants is a measure of the distance between graphs. Consequently, several metrics on the isomorphism classes of graphs have been proposed. See, for example, [1–9].

By a graph G we mean a simple one, i.e., without loops or multiple edges, with vertex set $V(G)$ and edge set $E(G)$. One way to obtain a metric is the following. Suppose $\stackrel{\mathscr{E}}{\rightarrow}$ denotes a symmetric nonreflexive binary relation on (the isomorphism classes of) graphs. Then we say that graph G can be transformed into graph H in k steps by $\mathscr{E}$ if there exists a sequence $G = G_0, G_1, G_2, \ldots, G_k = H$ of graphs such that $G_i \stackrel{\mathscr{E}}{\rightarrow} G_{i+1}$ for $0 \leqslant i \leqslant k-1$. The *distance* $\delta_{\mathscr{E}}(G,H)$ between G and H is defined as the minimum value of k such that G can be transformed into H in k steps by $\mathscr{E}$, if such a k exists; otherwise the distance is defined to be ∞. Thus, $\delta_{\mathscr{E}}$ is a metric on the class of graphs, though not all pairs of graphs need be a finite distance apart.

The three metrics we study here are derived from edge manipulation. Define $G \stackrel{\mathrm{EM}}{\rightarrow} G'$ if there exist $e_1 \in E(G)$ and $e_2 \in E(\bar{G})$ such that $G' = G - e_1 + e_2$.

* Corresponding author.

[1] The research of the first author was supported by ONR grant N00014-91-J-1022.

SSDI 0012-365X(95)00073-9

We call the associated metric δ_{EM} the *edge move distance*. It is convenient to talk of G' as being formed from G by *moving* the edge e_1.

Define $G \overset{\mathrm{ER}}{\rightarrow} G'$ if there exist $e_1 \in E(G)$ and $e_2 \in E(\bar{G})$ such that e_1 and e_2 have a vertex in common and $G' = G - e_1 + e_2$. We say that G is transformed into G' by an *edge rotation*, and call the associated metric δ_{ER} the *edge rotation distance*.

Define $G \overset{\mathrm{ES}}{\rightarrow} G'$ if there exist $e_1 = xy \in E(G)$ and $e_2 = xz \in E(\bar{G})$ such that y and z are adjacent and $G' = G - e_1 + e_2$. We say that G is transformed into G' by an *edge slide*, and call the associated metric δ_{ES} the *edge slide distance*. Thus an edge slide is a special type of edge rotation, which in turn is a special type of edge move. See Fig. 1 for illustration.

The edge rotation distance was introduced by Chartrand et al. [3] while the edge slide distance was defined by Johnson [5]. Though the name is from [2], the edge move distance was first defined by Baláž et al. [1] and Johnson [4]. (They allowed the graphs to have differing numbers of vertices and edges: this can be accommodated by allowing the removal or insertion of an edge or isolated vertex at unit cost.)

For the three metrics we study, for two graphs to be a finite distance apart they must have the same number of vertices and edges. This is sufficient for edge moves and rotations, but not for edge slides as they preserve connectivity.

It is immediate that for all graphs G and H:

$$\delta_{\mathrm{EM}}(G,H) \leqslant \delta_{\mathrm{ER}}(G,H) \leqslant \delta_{\mathrm{ES}}(G,H). \tag{1}$$

It has also been shown that (cf. [3]):

$$\delta_{\mathrm{ER}}(G,H) \leqslant 2\,\delta_{\mathrm{EM}}(G,H). \tag{2}$$

In contrast the ratio $\delta_{\mathrm{ES}}(G,H)/\delta_{\mathrm{ER}}(G,H)$ can be made arbitrarily large, and indeed infinite.

Another useful result is that there is an equivalent formulation for the edge move distance. A *common subgraph* of two graphs G and H is a graph F which is (isomorphic to) a subgraph of both G and H. If G and H have the same number of vertices, one can always extend F to have their number of vertices, i.e. F is a common spanning subgraph. The *greatest common subgraph size* of G and H, denoted $\mathrm{gcs}(G,H)$, is the maximum number of edges in a common subgraph of G and H. Then (cf. [2]):

$$\delta_{\mathrm{EM}}(G,H) = q - \mathrm{gcs}(G,H), \tag{3}$$

for G and H with q edges and n vertices.

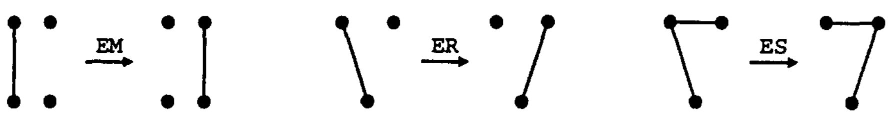

Fig. 1. An edge move, edge rotation and edge slide.

In this paper we further explore these three metrics. We look first at bounds on and relations between them. Then we consider their sensitivity to various graph operations. Finally, we look at the distance between arbitrary trees.

2. Bounds

In this section we look at bounds on and relationships between the three metrics. We start with two simple but useful bounds.

Theorem 1. *Let graphs G and H have n vertices and* $p\binom{n}{2}$ *edges. Then* $\delta_{\text{EM}}(G,H) \leqslant p(1-p)\binom{n}{2}$.

Proof. Consider a random bijection φ from $V(G)$ to $V(H)$. We are interested in the size of the common subgraph induced by φ. For any edge e in G, the probability that φ maps e to an edge in H is p. Thus, the expected number of edges preserved by φ is $p^2\binom{n}{2}$. Thus there is a φ which preserves at least this many, and hence $\text{gcs}(G,H) \geqslant p^2\binom{n}{2}$. Thus, $\delta_{\text{EM}}(G,H) \leqslant p\binom{n}{2} - p^2\binom{n}{2}$, as required. □

It is easy to show that $\delta_{\text{EM}}(G,H) \geqslant |\Delta(G) - \Delta(H)|$ (where $\Delta(G)$ denotes the maximum degree of G). But a more general bound for edge rotation distance is given by:

Theorem 2. *Let graph G have degree sequence* $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_n$ *and let graph H have degree sequence* $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_n$. *Then*

$$\delta_{\text{ER}}(G,H) \geqslant \tfrac{1}{2} \sum_i |d_i - e_i|.$$

Proof. An edge rotation increases the degree of one vertex and decreases the degree of one vertex. If one allows increasing or decreasing the degree of a vertex at $\frac{1}{2}$ cost, then the distance between the two degree sequences is the above bound. □

We conclude this section with a pair of results similar to Eq. 3. This requires some more concepts. We define a *labelled graph* as a graph G^* with n vertices and q edges where the vertices are labelled 1 up to n and the edges are labelled 1 up to q. For $1 \leqslant i \leqslant q$ we define $S_{G^*}(i)$ to be the pair of labels of endpoints of the edge labelled i; this pair is the *slot* occupied by the edge labelled i. We like to think of a labelled graph as a physical model in which the vertices are n numbered pegs on a board, there are $\binom{n}{2}$ slots, and q numbered wires each occupying one slot. Then an *edge move in a labelled graph* entails moving an edge from one slot to another, and possibly renumbering some edges. If no renumbering of the edges occurs we say that this edge move is *edge-label-preserving*.

A *labelling* of a graph G is a labelled graph G^* isomorphic to G. It follows that two graphs G and H are isomorphic iff there exist identical labellings G^* and H^* of the two. If G^* and H^* are labelled graphs on the same number of vertices and edges, we define for $1 \leqslant i \leqslant q$ the *slot-shift* of i by

$$SS_{G^*,H^*}(i) = 2 - |S_{G^*}(i) \cap S_{H^*}(i)|.$$

We will write $SS(i)$ if the labellings are clear from the context. An example is shown in Fig. 2. There $SS(1) = 2$, $SS(4) = 1$ and $SS(2) = SS(3) = 0$.

Theorem 3. *For all graphs G and H on n vertices and q edges*

$$\delta_{\mathrm{EM}}(G,H) = q - \mathrm{gcs}(G,H) = \min \sum_{i=1}^{q} \lceil SS_{G^*,H^*}(i)/2 \rceil,$$

where the minimum is taken over all labellings G^* and H^* of G and H.

Proof. Let $SSS(G^*,H^*) = \sum_{i=1}^{q} \lceil SS_{G^*,H^*}(i)/2 \rceil$. We show that: $\min SSS(G^*,H^*) = q - \mathrm{gcs}(G,H)$.

Let F_G and F_H be isomorphic spanning subgraphs of G and H with c edges. Then there exists a bijection $\varphi : V(G) \to V(H)$ that maps $E(F_G)$ onto $E(F_H)$. Let G^* be any labelling of G. Define a labelling H^* of H such that $\varphi(v)$ receives the same label as v ($v \in V(G)$), and if $\{v,w\}$ is an edge of F_G then it and $\{\varphi(v),\varphi(w)\}$ receive the same label. For this pair of labellings, the value of SSS is at most $q - c$. It follows that: $\min SSS(G^*,H^*) \leqslant q - \mathrm{gcs}(G,H)$.

For the reverse inequality, consider any labellings G^* and H^*. Then the collection of edges such that $SS(i) = 0$ induce identical subgraphs of G^* and H^*. If follows that $SSS(G^*,H^*) \geqslant q - \mathrm{gcs}(G,H)$. □

Theorem 4. *For all graphs G and H on n vertices and q edges*

$$\delta_{\mathrm{ER}}(G,H) = \min \sum_{i=1}^{q} SS_{G^*,H^*}(i),$$

where the minimum is taken over all labellings G^* and H^* of G and H.

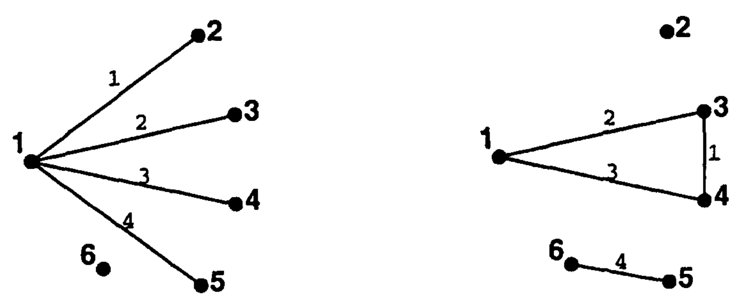

Fig. 2. Two labelled graphs.

Proof. Let $SSS(G^*,H^*) = \sum_{i=1}^{q} SS_{G^*,H^*}(i)$ for arbitrary G^* and H^*. Let G^* and H^* be specific labellings of G and H. We show first that $SSS(G^*,H^*)$ edge rotations suffice to transform G^* into H^*. Note that the summand in SSS is 0 if the edge labelled i is in the same slot in G^* and H^*, 1 if the slots overlap in one vertex, and 2 if the slots are disjoint. It is immediate that one could transform G^* into H^* by $SSS(G^*,H^*)$ edge rotations if one were allowed multiple edges at intermediate steps.

The problem is to effect this transformation without multigraphs. We show this by induction on the number of labels i such that $SS_{G^*,H^*}(i) > 0$. In particular, we show how to transform G^* into a new labelled graph G'^* such that the decrease in the value of SSS (i.e., $SSS(G^*,H^*) - SSS(G'^*,H^*)$) is at least the number of edge rotations used.

Let i be the label of any edge such that $SS(i) > 0$. If $SS(i) = 1$ and the slot $S_{H^*}(i)$ is unoccupied in G^*, then rotate the edge. For 1 edge rotation the value of SSS is decreased by 1. If $S_{H^*}(i)$ is occupied by an edge labelled j (say), then swop the labels i and j. Since H^* is a graph, the edge labelled j was in the wrong slot for H^* anyway. So this action, which does not cost any edge rotations, decreases $SS(i)$ to 0 and increases $SS(j)$ by at most 1, so that SSS does not increase. Now suppose that $SS(i) = 2$ and $S_{G^*}(i) = \{u,v\}$ and $S_{H^*}(i) = \{x,y\}$. If slots $\{u,x\}$ and $\{x,y\}$ are unoccupied in G^* then rotate edge i from $\{u,v\}$ to $\{u,x\}$ to $\{x,y\}$. For 2 edge rotations we reduce SSS by 2. If slot $\{x,y\}$ is unoccupied but $\{u,x\}$ is occupied by an edge labelled j (say), then rotate edge j from $\{u,x\}$ to $\{x,y\}$, rotate edge i from $\{u,v\}$ to $\{u,x\}$ and swop labels i and j. For 2 edge rotations we reduce SSS by 2. If slot $\{x,y\}$ is occupied by an edge labelled k (say), swop labels i and k. Since H is a graph, edge k is in the wrong slot anyway. So for 0 edge rotations we reduce $SS(i)$ by 2 and increase $SS(k)$ by at most 1 for a net decrease in SSS. In every case we reduce the number of edges in incorrect slots. It follows that $\delta_{\mathrm{ER}}(G,H) \leqslant SSS(G^*,H^*)$.

For the reverse inequality, suppose $\delta_{\mathrm{ER}}(G,H) = k$. Then there exists a sequence $G = G_0, G_1, \ldots G_k = H$ of graphs, each one obtained from the previous one by one edge rotation. Choose any labelling of G_0. Then we can choose labellings of $G_1, G_2, \ldots, G_k$ such that G_{i+1}^* is obtained from G_i^* by a single edge-label-preserving edge rotation.

Now, G_0^* and G_k^* are labellings of G and H. Since the process is edge-label-preserving, the edge labelled i must be rotated at least $SS(i)$ times in going from G_0^* to G_k^*. It follows that k is at least $SSS(G_0^*, G_k^*)$ which shows that $\delta_{\mathrm{ER}} \geqslant \min SSS$, as required. □

The pair of labellings given in Fig. 2 are ones that minimize the value of SSS in the above two theorems.

Corollary 1. *The edge move and edge rotation distances between two graphs remain the same even if one allows multigraphs at intermediate steps.*

Proof. This follows for edge rotation distance by the same argument used in the final two paragraphs of the proof of Theorem 4: If $G = G_0, G_1, \ldots G_k = H$ is a minimum sequence of multigraphs, each one obtained from the previous one by one edge rotation, then there is a labelling of the multigraphs such that the sequence of edge rotations is edge-label-preserving, and thus k is at least $SSS(G_0^*, G_k^*)$. The corresponding inequality for edge move distance can be proved similarly. □

2.1. Adding a new vertex

We consider the operations of adding an isolated or dominating vertex. By the common subgraph formulation, it is immediate that both these operations preserve the edge move distance. The same is true of edge rotation distance:

Theorem 5. $\delta_{\mathrm{ER}}(G \cup K_1, H \cup K_1) = \delta_{\mathrm{ER}}(G, H)$.

Proof. The inequality $\delta_{\mathrm{ER}}(G \cup K_1, H \cup K_1) \leqslant \delta_{\mathrm{ER}}(G, H)$ is clear. We prove the reverse inequality.

Consider the labellings $(G \cup K_1)^*$ and $(H \cup K_1)^*$ for which the sum SSS in Theorem 4 is minimized. Say the (designated) isolated vertex is labelled x in $(G \cup K_1)^*$ and labelled y in $(H \cup K_1)^*$. (Fig. 2 shows that it is possible for the isolated vertices to have different labels in an optimal pair of labellings.) Construct a new labelling $(H \cup K_1)^{**}$ by taking the original labelling of $H \cup K_1$ and swopping the labels x and y. (If $x = y$ nothing changes.) We claim that the value of SSS stays the same. In fact, the value of $SS(i)$ is the same for the two labellings, unless the edge with label i is incident with y in $(G \cup K_1)^*$ and with x in $(H \cup K_1)^*$, in which case $SS(i)$ decreases by 1 when we relabel. By the optimality of the original labellings the latter case cannot occur; so the new labelling is optimal too.

By deleting the vertex labelled x from both graphs, we obtain labellings G^* and H^{**} of G and H. It follows that $\delta_{\mathrm{ER}}(G, H) \leqslant SSS((G \cup K_1)^*, (H \cup K_1)^*) = \delta_{\mathrm{ER}}(G \cup K_1, H \cup K_1)$. □

Since the edge rotation distance between two graphs G and H is the same as the edge rotation distance between their complements $\bar{G}$ and $\bar{H}$ (cf. [3]), it follows that $\delta_{\mathrm{ER}}(G + K_1, H + K_1) = \delta_{\mathrm{ER}}(\bar{G} \cup K_1, \bar{H} \cup K_1) = \delta_{\mathrm{ER}}(\bar{G}, \bar{H}) = \delta_{\mathrm{ER}}(G, H)$ and, thus,

$$\delta_{\mathrm{ER}}(G \cup K_1, H \cup K_1) = \delta_{\mathrm{ER}}(G, H) = \delta_{\mathrm{ER}}(G + K_1, H + K_1).$$

We now consider the edge slide distance. Adding an isolated vertex clearly does not alter the edge slide distance. On the other hand, joining a vertex can considerably reduce the slide distance:

Theorem 6. $\delta_{\mathrm{ES}}(G + K_1, H + K_1) \leqslant 2\, \delta_{\mathrm{ER}}(G, H)$.

Proof. It suffices to show that if a graph G can be transformed into a graph G' by an edge rotation, then $G + K_1$ can be transformed into $G' + K_1$ by two edge slides. Take a labelling of $G + K_1$ where the designated dominating vertex is labelled $n + 1$. Suppose the edge rotation from G to G' involves moving an edge from the slot $\{x,y\}$ to the slot $\{x,z\}$. Then in the labelling of $G+K_1$ one can effect this by sliding an edge from $\{n+1,x\}$ to $\{x,z\}$ and then sliding an edge from $\{x,y\}$ to $\{n+1,x\}$. □

As a consequence of the above two results we obtain:

Theorem 7. *If G and H are connected graphs on $n \geqslant 2$ vertices then $\delta_{ES}(G,H) \leqslant 2\,\delta_{ER}(G,H)+6n-12$.*

Proof. Let v be a vertex of G and let T_v be a spanning tree of G including all the edges incident with v. Let $G' = G - E(T_v)$, and define G'' to be obtained from G' by adding edges from v to all other vertices of G. (Note that G and G'' have the same number of edges.) For H define a vertex w, a spanning tree U_w, a pruned graph H' and an extended graph H'' similarly.

We transform G to H as follows: first we transform G into G'', then G'' into H'', and then H'' into H. The first stage is accomplished by transforming T_v to a star with center v. This takes at most $n-2$ edge slides. (At each step find an occupied slot $\{w,x\}$ where $\{v,w\}$ is occupied but $\{v,x\}$ is not and slide that edge to the slot $\{v,x\}$.) The third stage is accomplished by transforming a star with center w to U_w. Hence, $\delta_{ES}(G,H) \leqslant \delta_{ES}(G'',H'')+2n-4$.

It is not hard to see that G' can be transformed into H' by at most $\delta_{ER}(G,H)$ edge rotations followed by $n-2$ edge moves. Thus, $\delta_{ER}(G'-v,H'-w) \leqslant \delta_{ER}(G,H)+2(n-2)$. By the above two theorems, this means that $\delta_{ES}(G'',H'') \leqslant 2\,\delta_{ER}(G,H)+4(n-2)$. □

2.2. Maximum distance

Theorem 8. *The maximum distance between two graphs on n vertices under edge move distance is $n^2/8 - \Theta(n)$.*

Proof. Theorem 1 shows that the distance between two graphs is at most $\binom{n}{2}/4$. Near-equality is obtained, for example, by $G = K(l,l) \cup K_1$ and $H = K_l \cup K_{l+1}$ where $n = 2l+1$. □

(A positive function $f(n)$ is $\Theta(n)$ if there exist positive constants c_1 and c_2 such that $c_1 n \leqslant f(n) \leqslant c_2 n$ for all n.)

Using Eq. 2 and Theorem 7, we see that the maximum distance between graphs under edge rotations is at most $n^2/4$, and at most $n^2/2+6n$ under edge slide.

However, the biggest distance under edge rotation we know of is about $4n^2/27$. By the degree-sequence bound (Theorem 2), (at least) this is achieved by G an approximately $4n/9$-regular graph, and $H = K_{2n/3} \cup (n/3)K_1$.

3. Graph operations

We have already looked at what happens when one adds a dominating or isolated vertex. So now we consider what happens for other operations.

For edge slides it is immediate that the distance doubles when one considers $2G$ and $2H$ rather than G and H. One might expect the same to be true of the other two metrics. But it is not:

Theorem 9. *There exist graphs G and H such that $\delta(2G,2H) < 2\delta(G,H)$ under edge move and edge rotation distance.*

Proof. Let F be a highly connected graph with distinguishable vertices x and y.

Let G be the disjoint union of four copies of F, $F_1,\ldots,F_4$, and four isolated vertices, together with the four edges connecting:

$$F_1(x) - F_2(x), \quad F_3(x) - F_4(x), \quad F_1(y) - F_3(y) \quad \text{and} \quad F_2(y) - F_4(y),$$

where (for example) $F_1(x)$ denotes the vertex corresponding to x in F_1.

Let H be the disjoint union of four copies of F, $F'_1,\ldots,F'_4$, and $2K_2$, together with the two edges connecting:

$$F'_1(x) - F'_2(x) \quad \text{and} \quad F'_3(y) - F'_4(y).$$

See Fig. 3.

By the connectivity of F, the copies of F must be preserved in any transformation of G to H which uses only a few edge operations, and similarly with $2G$ to $2H$. So $\delta_{\mathrm{EM}}(G,H) = 3$ and $\delta_{\mathrm{ER}}(G,H) = 5$. But one can improve on this procedure when taking two copies, and show that $\delta_{\mathrm{EM}}(2G,2H) = 4$. (See Fig. 3.) This shows that $\delta_{\mathrm{ER}}(2G,2H) = 8$. □

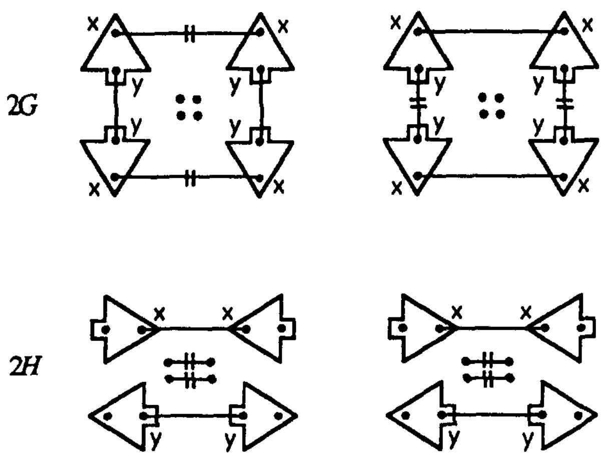

Fig. 3. Moving the notched edges transforms $2G$ into $2H$.

For arbitrary unions note that $G_1 \cup G_2$ and $H_1 \cup H_2$ may be isomorphic even though no G_i and H_j are isomorphic.

3.1. Subdivisions

Let $S(G)$ denote G with every edge subdivided once. It is useful to think of $S(G)$ defined so that $V(S(G)) = V(G) \cup E(G)$. Since an element of $E(G)$ plays different roles in G and $S(G)$, we will refer to a subdivision as having nodes and lines rather than vertices and edges. If G has n vertices and q edges then $S(G)$ has $n+q$ nodes and $2q$ lines.

Theorem 10. *For all graphs G and H, $\delta_{\mathrm{ER}}(S(G),S(H)) \leqslant \delta_{\mathrm{ER}}(G,H)$.*

Proof. We need only show that a rotation $xy \to xz$ in G can be simulated by a rotation in $S(G)$. But let e be the node subdividing xy: then one may rotate $ey \to ez$, as required. □

As a corollary it follows that $\delta_{\mathrm{EM}}(S(G),S(H)) \leqslant 2\,\delta_{\mathrm{EM}}(G,H)$. Equality is possible: take G to be a path and H a tree which is formed by taking two disjoint paths and adding one edge joining an interior vertex on each. (To transform G to H takes one move, but to transform $S(G)$ to $S(H)$ takes two moves.)

There is a partial converse:

Theorem 11. *Let G and H be graphs with n vertices and q edges. Let φ be a bijection between $V(S(G))$ and $V(S(H))$ that preserves as many lines as possible, viz.* $\mathrm{gcs}(S(G),S(H))$.
If *φ maps $V(G)$ to $V(H)$* **then** *$\delta_{\mathrm{ER}}(G,H) \leqslant \delta_{\mathrm{EM}}(S(G),S(H))$, so that in this case*

$$\delta_{\mathrm{EM}}(G,H) \leqslant \delta_{\mathrm{EM}}(S(G),S(H)) = \delta_{\mathrm{ER}}(S(G),S(H)) = \delta_{\mathrm{ER}}(G,H).$$

Proof. Take any labelling $S(G)^*$ of $S(G)$. Define a labelling $S(H)^*$ of $S(H)$ such that $\varphi(v)$ receives the same label as v for every node v of $S(G)$, and if $\{v,w\}$ is a line of $S(G)$ and $\{\varphi(v),\varphi(w)\}$ a line of $S(H)$ then the two lines receive the same label. Note that $\sum_{k=1}^{2q} \lceil SS(k)/2 \rceil = 2q - \mathrm{gcs}(S(G),S(H)) = \delta_{\mathrm{EM}}(S(G),S(H))$ for this pair of labellings.

Assume φ maps $V(G)$ to $V(H)$. It thus maps the remaining nodes of $S(G)$, viz. $E(G)$, to the corresponding nodes of $S(H)$, viz. $E(H)$. In particular, the set of labels received by $V(G)$ is the same as the set of labels received by $V(H)$, and similarly with $E(G)$ and $E(H)$. Define labellings G^* and H^* by giving the vertices and edges of G and H the same labels they received as nodes of $S(G)$ and $S(H)$. (Let us extend the definition of a labelled graph to allow this.)

Consider an element of $E(G)$ labelled e. Let $S_{G^*}(e) = \{x,y\}$ and $S_{H^*}(e) = \{a,b\}$. This means that the neighbors of the node labelled e are labelled x and y in $S(G)$ and are labelled a and b in $S(H)$. Then let i_e and j_e be the labels of the lines of $S(G)$ which join the node labelled e to the nodes labelled x and y, respectively.

If $SS_{S(G)^*,S(H)^*}(i_e) = 0$ then the line labelled i_e is incident with the node labelled e in $S(H)^*$ and so $x \in \{a,b\}$. Further, if $SS_{S(G)^*,S(H)^*}(i_e)$ and $SS_{S(G)^*,S(H)^*}(j_e)$ are both zero then $\{x,y\} = \{a,b\}$. Hence,

$$SS_{G^*,*}(e) \leqslant \lceil SS_{S(G)^*,H)^*}(i_e)/2\rceil + \lceil SS_{S(G)^*,H)^*}(j_e)/2\rceil.$$

If we sum the left-hand expression over all e we obtain an upper bound on $\delta_{\mathrm{ER}}(G,H)$ by Theorem 4. If we sum the right-hand side expression over all e we obtain $\lceil SS(k)/2\rceil$ exactly once for each line-label. So that sum is $\delta_{\mathrm{EM}}(S(G),S(H))$. The result follows. □

Surprisingly, perhaps, φ need not map $V(G)$ to $V(H)$. Indeed:

Theorem 12. *There exist G and H such that $\delta_{\mathrm{EM}}(S(G),S(H)) < \delta_{\mathrm{EM}}(G,H)$.*

Proof. Let F be a highly connected graph with four distinguished vertices v, w, x and y. Form G and H by taking F and adding four vertices and four edges: in G add a path of length two between v and w and a disjoint P_3 as a component; in H add a path of length three between x and y and a disjoint P_2 as a component.

It is easy to see that the greatest common subgraph of G and H contains all of F, but no added edge incident with v, w, x or y, and indeed only one added edge. Thus, $\delta_{\mathrm{EM}}(G,H) = 3$.

However, $S(G)$ and $S(H)$, less the two spliced added edges incident with v and w, or x and y as the case may be, are isomorphic. Thus $\delta_{\mathrm{EM}}(S(G),S(H)) = 2$. □

It remains unresolved whether $\delta_{\mathrm{ER}}(S(G),S(H)) = \delta_{\mathrm{ER}}(G,H)$ always.

4. Trees

The previous results on trees gave the distance to the path P_n or star S_n:

Proposition 1 (Benadé [2]; Zelinka [8,9]). *For all trees T on n vertices,*

(1) $\delta_{\mathrm{EM}}(T,S_n) = \delta_{\mathrm{ER}}(T,S_n) = \delta_{\mathrm{ES}}(T,S_n) = n - 1 - \Delta(T)$;
(2) $\delta_{\mathrm{ES}}(T,P_n) = n - 1 - \mathrm{diam}(T)$;
(3) $\delta_{\mathrm{ER}}(T,P_n) = \mathrm{end}(T) - 2$;

where $\Delta(T)$, $\mathrm{diam}(T)$ and $\mathrm{end}(T)$ denote the maximum degree, diameter, and number of end-vertices, of T.

Note that both edge rotation results are examples of equality in the degree-sequence bound. A succinct formula for $\delta_{\mathrm{EM}}(T,P_n)$ is not known.

But what about distance between arbitrary trees? Any two trees on $n \geqslant 3$ vertices are at most $n-3$ edge moves apart since they contain P_3 as a common subgraph. But the following is true.

Theorem 13. *For all trees T and U on $n \geqslant 3$ vertices, $\delta_{\mathrm{ER}}(T,U) \leqslant n-3$.*

Proof. The proof is by induction on n. The statement is true for $n=3$ as then $T \cong U \cong P_3$; so let $n \geqslant 4$. Let t be an end-vertex of T with neighbor t', and u an end-vertex of U with neighbor u'. By the induction hypothesis, $\delta_{\mathrm{ER}}(T-t, U-u) \leqslant n-4$. Then the edge tt' can be rotated to tu' to complete the transformation. □

For edge slides the situation is unresolved. Clearly, the edge slide distance between two trees is at most $\delta_{\mathrm{ES}}(T,S_n) + \delta_{\mathrm{ES}}(S_n,U) = 2(n-3)$. This can easily be improved, but still only to $2n - \mathrm{o}(n)$. (A function $f(n)$ is $\mathrm{o}(n)$ if $\lim_{n\to\infty} f(n)/n = 0$.)

This suggests that one should look for "central" trees. But there is no good central tree even for edge rotation distance.

Theorem 14. *For all trees T on n vertices there exists a tree U on n vertices such that $\delta_{\mathrm{ER}}(T,U) \geqslant n - \mathrm{o}(n)$.*

We consider the family of k-ary trees and use the degree-sequence bound (Theorem 2) to show that no tree is close to all such trees. The proof is a straightforward result on the approximation of the family of "degree-functions," and could be omitted without loss of continuity.

For k and n such that $k|n-2$, let $T_n(k)$ be a tree on n vertices with degree set $\{1, k+1\}$. Then define the real function $f_{k,n}(i)$ as k on $i \in (0,(n-2)/k]$ and 0 otherwise. On the set $\{1,2,\ldots,n\}$, this function gives 1 less than the degree-sequence of $T_n(k)$. We show that there is no real function close to all of $\{f_{k,n}\}_k$.

Lemma 1. *Let $m = 2^{r2^r}$ and $n = m+2$. Consider the ℓ_1-norm. Let*

$$\rho_n = \min_{f_n} \max_{k} ||f_{k,n} - f_n||,$$

where $||f_n|| = m$. Then $\rho_n \geqslant 2m - 4m/2^r$.

Proof. Suppose one takes only an increasing subset $\{a_1,\ldots,a_s\}$ of the possible values of k. Let f be the best approximation to this collection with $||f|| = m$.

It is immediate that f is decreasing. One may assume that it is piecewise linear, with value b_i on $(m/a_{i+1}, m/a_i]$ (where $m/a_{s+1} = 0$). By the area constraint, $b_i \leqslant a_i$. Indeed,

$$||f_{a_i,n} - f|| \geqslant 2(a_i - b_i)(m/a_i - m/a_{i+1}).$$

But $\sum_i b_i(m/a_i - m/a_{i+1}) = ||f|| = m$. Thus,

$$\sum_{i=1}^{s} ||f_{a_i,n} - f|| \geqslant 2(ms - m - m\sum_i a_i/a_{i+1}).$$

Now set $s = 2^r + 1$ and $a_i = 2^{r(i-1)}$ $(i = 1,\ldots,s)$. Then

$$\max_i \left(||f_{a_i,n} - f||\right) \geqslant 2(m - m/s - m/2^r)$$

as required. □

Other values of n can be accommodated. Thus for all trees T on n vertices there exists a tree U such that $\delta_{\mathrm{ER}}(T,U) \geqslant n - 2n \log\log n/\log n -$ smaller order terms.

References

[1] V. Baláž, J. Koča, V. Kvasnička, and M. Sekanina, A measure for graphs, Časopis Pěst. Mat. 111 (1986) 431–433.

[2] G. Benadé, W. Goddard, T.A. McKee and P.A. Winter, On distances between isomorphism classes of graphs, Math. Bohemica 116 (1991) 160–169.

[3] G. Chartrand, F. Saba and H. Zou, Edge rotations and distance between graphs, Časopis Pěst. Mat. 110 (1985) 87–91.

[4] M. Johnson, Relating metrics, lines and variables defined on graphs to problems in medicinal chemistry, in: (Y. Alavi et al. eds), Graph Theory with Applications to Algorithms & Computer Science (Wiley, New York, 1985) 457–470.

[5] M. Johnson, An ordering of some metrics defined on the space of graphs, Czech. Math. J. 37 (1987) 75–85.

[6] B. Zelinka, On a certain distance between isomorphism classes of graphs, Časopis Pěst. Mat. 100 (1975) 371–373.

[7] B. Zelinka, A distance between isomorphism classes of trees, Czech. Math. J. 33 (1983) 126–130.

[8] B. Zelinka, Comparison of various distances between isomorphism classes of graphs, Časopis Pěst. Mat. 110 (1985) 289–293.

[9] B. Zelinka, Contraction distance between isomorphism classes of graphs, Časopis Pěst. Mat. 115 (1990) 211–216.