Časopis pro pěstování matematiky

Vladimír Baláž; Vladimír Kvasnička; Jiří Pospíchal

Dual approach to edge distance between graphs

Časopis pro pěstování matematiky, Vol. 114 (1989), No. 2, 155--159

Persistent URL: `http://dml.cz/dmlcz/108715`

DUAL APPROACH TO EDGE DISTANCE BETWEEN GRAPHS*)

VLADIMÍR BALÁŽ, VLADIMÍR KVASNIČKA, JIŘÍ POSPÍCHAL, Bratislava

(Received January 26, 1987)

Summary. Two kinds of edge distances between graphs are defined. They are based on the notion of maximal common subgraph and minimal common supergraph. These distances form metrics for isomorphism classes of graphs and, moreover, they are fully identical.

Keywords: Graph metric, maximal common subgraph, minimal common supergraph, graph distance.

AMS Classification: 05C99, 68E10.

1. In our recent communication [1], and independently [6] we have introduced a metric for graphs conceptually based on the notion of maximal common subgraph (MCS) of two graphs. This kind of a metric is very well suited for mathematical modelling of organic chemistry [2], the evaluated distance between two graphs (molecules) being closely related to the number of elementary reaction steps transforming an educt molecule into a product one. The purpose of this communication is to reformulate the above mentioned metric in terms of the notion dual to the MCS; in particular, the so-called minimal common supergraph (mcs). We shall compare these two possibilities how to define the edge metric. Recently, other metrics for graphs were compared with our edge metric by Zelinka [3].

2. A graph (directed or undirected) G is composed of a non-empty finite vertex set $V(G)$ and an edge set $E(G)$. A subgraph G' of a graph G is a graph obtained from G by deleting subsets (which may be empty) of its vertex and edges sets, notation $G' \subseteq G$; G' is contained in G which is a supergraph of G', $G \supseteq G'$. Two graphs G_1 and G_2 are isomorphic, $G_1 \sim G_2$, if there exists a 1-1 correspondence between the vertices of G_1 and G_2 such that the adjacent pairs of vertices in G_1 are 1-1 mapped only to adjacent pairs of vertices in G_2. A common subgraph (supergraph) of two graphs G_1 and G_2 consists of a subgraph $G_1' \subseteq G_1$ (supergraph $G_1' \supseteq G_1$) and a subgraph $G_2' \subseteq G_2$ (supergraph $G_2' \supseteq G_2$) such that $G_1' \cong G_2'$. A maximal common subgraph (minimal common supergraph) of two graphs G_1 and G_2 is a common subgraph (supergraph) which contains the largest (smallest) possible number of edges; it will be denoted by $G_1 \cap G_2$ ($G_1 \cup G_2$). We emphasize that the concept

*) Part X in the series Mathematical Model of Organic Chemistry.

of MCS/mcs can easily be generalized by induction to more than two graphs. Let G be a graph with n vertices, i.e. $|V(G)| = n$. A complement $\bar{G}$ of the graph G with respect to the complete graph K_n (where $G \subseteq K_n$) is the graph obtained from K_n by deleting all edges of G.

3. Let us have a pair of graphs G and H. We assign to them two nonnegative integer valued functions. The function $g(G, H)$ expresses the number of edges that can be deleted from G so that the resulting graph is isomorphic to their MCS denoted by $G \cap H$. Similarly, the function $h(G, H)$ expresses the number of edges that can be added to G so that the resulting graph is isomorphic to their mcs denoted by $G \cup H$. These verbal definitions of functions g and h can be formally presented as

$$g(G, H) = |E(G)| - |E(G \cap H)| , \tag{1a}$$

$$h(G, H) = |E(G \cup H)| - |E(G)| , \tag{1b}$$

where $|E(X)|$ is the cardinality (number of elements) of the edge set $E(X)$ of a graph X.

Lemma 1. *The function g satisfies*

(2a) (i) *if* $G \supseteq H$, *then* $g(G, H) = |E(G)| - |E(H)|$;

(2b) (ii) $g(G, H) = g(G, G \cap H)$;

(2c) (iii) *if* $G_1 \subseteq G_2$, *then* $g(G_1, H) \leqq g(G_2, H)$;

(2d) (iv) *if* $H_1 \supseteq H_2$, *then* $g(G, H_1) \leqq g(G, H_2)$;

(2e) (v) *if* $G \supseteq H \supseteq T$, *then* $g(G, H) + g(H, T) = g(G, T)$;

(2f) (vi) $g(G, H) \geqq g(G \cap T, H)$ *for any graph* T.

All these properties are obvious, they have very simple and heuristic set-theoretic interpretation and meaning.

Lemma 2. *The function h satisfies*

(3a) (i) *if* $G \subseteq H$, *then* $h(G, H) = |E(H)| - |E(G)|$;

(3b) (ii) $h(G, H) = h(G, G \cup H)$;

(3c) (iii) *if* $G_1 \supseteq G_2$, *then* $h(G_1, H) \leqq h(G_2, H)$;

(3d) (iv) *if* $H_1 \subseteq H_2$, *then* $h(G, H_1) \leqq h(G, H_2)$;

(3e) (v) *if* $G \subseteq H \subseteq T$, *then* $h(G, H) + h(H, T) = h(G, T)$;

(3f) (vi) $h(G, H) \geqq h(G \cup T, H)$ *for any graph* T.

Similarly as above, all these properties are obvious with a very simple set-theoretic interpretation.

4. We shall define two kinds of edge distances between two graphs employing the notion of MCS and mcs, respectively.

Definition 1. The edge distance d between two graphs G and H is the number

$$d(G, H) = g(G, H) + g(H, G) + \big| |V(G)| - |V(H)| \big| . \tag{4}$$

Definition 2. The edge distance D between two graphs G and H is the number

$$D(G, H) = h(G, H) + h(H, G) + \big| |V(G)| - |V(H)| \big| . \tag{5}$$

Introducing the relations (1a–b) into the right-hand sides of (4) and (5), respectively, we arrive at alternative definitions of the edge distances between two graphs,

$$d(G, H) = |E(G)| + |E(H)| - 2|E(G \cap H)| + \big| |V(G)| - |V(H)| \big| , \tag{6a}$$

$$D(G, H) = -|E(G)| - |E(H)| + 2|E(G \cup H)| + \big| |V(G)| - |V(H)| \big| , \tag{6b}$$

where the first relation (6a) coincides with our previous definition of the edge distance from [1].

Theorem 1. *The edge distance d forms a metric for classes of mutually isomorphic graphs, the following three fundamenal relations being satisfied:*

(7a) (i) *Nonnegativeness*

$$d(G, H) \geqq 0 \quad (=0 \text{ iff } G \cong H) .$$

(7b) (ii) *Symmetry*

$$d(G, H) = d(H, G) .$$

(7c) (iii) *Triangle inequality*

$$d(G, H) + d(H, T) \geqq d(G, T) .$$

Theorem 2. *The edge distance D forms a metric for classes of mutually isomorphic graphs, the following three fundamental relations being satisfied:*

(8a) (*i*) *Nonnegativeness*

$$D(G, H) \geqq 0 \quad (=0 \text{ iff } G \sim H) .$$

(8b) (ii) *Symmetry*

$$D(G, H) = D(H, G) .$$

(8c) (iii) *Triangle inequality*

$$D(G, H) + D(H, T) \geqq D(G, T) .$$

In order to prove these two theorems it is fully sufficient to focus our attention only on the triangle inequalities, since the nonnegativeness and symmetry immediately follow from the definitions of the distances d and D. Furthermore, the „vertex" terms in (7c) and (8c) can be ignored, since the inequality $\big| |V(G)| -$

$- \big||V(H)|\big| + \big||V(H)| - |V(T)|\big| \geqq \big||V(G)| - |V(T)|\big|$ immediately follows from the well-known inequality $|a - b| + |b - c| \geqq |a - c|$.
Hence, we have to verify the inequalities

$$g(G, H) + g(H, G) + g(H, T) + g(T, H) \geqq g(G, T) + g(T, G),$$
$$h(G, H) + h(H, G) + h(H, T) + h(T, H) \geqq h(G, T) + h(T, G).$$

They follow from

(9a) $$g(G, H) + g(H, T) \geqq g(G, T),$$

(9b) $$g(H, G) + g(T, H) \geqq g(T, G),$$

and

(10a) $$h(G, H) + h(H, T) \geqq h(G, T),$$

(10b) $$h(H, G) + h(T, H) \geqq h(T, G).$$

For instance, let us prove the inequality (10a):

$$\begin{aligned} g(G, H) + g(H, T) &\geqq g(G, G \cap H) + g(G \cap H, T) = \\ &= g(G, G \cap H) + g(G \cap H,\ G \cap H \cap T) = \\ &= g(G, G \cap H \cap T) \geqq g(G, G \cap T) = \\ &= g(G, T), \end{aligned}$$

where the properties (2b–f) are used. The other inequalities (9b) and (10a–b) are proved in a completely analogous way.

Theorem 3. *The distances d and D are identical, i.e., for an arbitrary pair of graphs G and H we have*

(11) $$d(G, H) = D(G, H).$$

Proof. This theorem is a consequence of the identity

$$g(G, H) = h(H, G).$$

Using the definitions (1a–b) we get

$$|E(G)| + |E(H)| = |E(G \cup H)| + |E(G \cap H)|,$$

which is obviously satisfied.

Theorem 4. *Let H and G be two graphs with the same number of vertices. Then* (cf. [4, 5])

(12a) $$d(G, H) = d(\bar{G}, \bar{H}),$$

(12b) $$D(G, H) = D(\bar{G}, \bar{H}),$$

where $\bar{G}$ and $\bar{H}$ are their complements.

Proof. It is fully sufficient to realize that $\overline{G \cup H} = \bar{G} \cap \bar{H}$ and $\overline{G \cap H} = \bar{G} \cup \bar{H}$, hence and $m = |E(K_n)|$

$$\begin{aligned} d(\bar{G}, \bar{H}) &= |E(\bar{G})| + |E(\bar{H})| - 2|E(\bar{G} \cap \bar{H})| = \\ &= |E(\bar{G})| + |E(\bar{H})| - 2|E(\overline{G \cup H})| = \\ &= m - |E(G)| + m - |E(H)| - 2(m - |E(G)| - |E(H)| + |E(G \cap H)|) = \\ &= |E(G)| + |E(H)| - 2|E(G \cap H)|, \end{aligned}$$

which was to be proved. The other identity follows by (11).

References

[1] *V. Baláž, J. Koča, V. Kvasnička, M. Sekanina:* A metric for graphs, Časopis pěst. mat. *111* (1986), 431—433.

[2] *V. Kvasnička, M. Kratochvíl, J. Koča:* Reaction graphs, Collect. Czech. Chem. Commun. *48* (1983), 2284—2304.

[3] *B. Zelinka:* Edge-distances between isomorphism classes of graphs, Časopis pěst. mat. (to appear).

[4] *B. Zelinka:* On a certain distance between isomorphism classes of graphs, Časopis pěst. mat. *100* (1975), 371—373.

[5] *G. Chartrand, F. Saba, H. B. Zou:* Edge rotations and distance between graphs, Časopis pěst. mat. *110* (1985), 87—91.

[6] *M. Johnson:* Relating metrics, lines and variables defined on the space of graphs, in: Graph theory and its applications to algorithms and computer science (1985), 457—470.

Souhrn

DUÁLNÍ PŘÍSTUP K HRANOVÉ VZDÁLENOSTI MEZI GRAFY

VLADIMÍR BALÁŽ, VLADIMÍR KVASNIČKA, JIŘÍ POSPÍCHAL

V práci byly definovány dva typy hranových vzdáleností mezi grafy, založené na pojmech „maximální společný podgraf" a „minimální společný nadgraf". Tyto vzdálenosti tvoří metriky pro isomorfní třídy grafů a navíc jsou plně identické.

Резюме

ДВОЙСТВЕННОЙ ПОДХОД К РЕБЕРНОМУ РАССТОЯНИЮ МЕЖДУ ГРАФАМИ

VLADIMÍR BALÁŽ, VLADIMÍR KVASNIČKA, JIŘÍ POSPÍCHAL

В работе определены два типа реберных расстояний между графами, основанные на понятиях ,,максимальный общий подграф" и ,,минимальный общий надграф" соответственно эти расстояния порождают метрики для классов изоморфизма графов и полностью совпадают.

Authors' address: Katedra matematiky, CHTF SVŠT, Radlinského 9, 812 37 Bratislava.