A graph distance metric combining maximum common subgraph and minimum common supergraph

Mirtha-Lina Fernández [a,1], Gabriel Valiente [b,2]

[a] *Computer Science Department*
University of Oriente
90900 Santiago de Cuba, Cuba
mirtha@csd.uo.edu.cu

[b] *Department of Software*
Technical University of Catalonia
E-08034 Barcelona, Spain
valiente@lsi.upc.es

Abstract

The relationship between two important problems in pattern recognition using attributed relational graphs, the maximum common subgraph and the minimum common supergraph of two graphs, is established by means of simple constructions, which allow to obtain the maximum common subgraph from the minimum common supergraph, and vice versa. On this basis, a new graph distance metric is proposed for measuring similarities between objects represented by attributed relational graphs. The proposed metric can be computed by a straightforward extension of any algorithm that implements error-correcting graph matching, when run under an appropriate cost function, and the extension only takes time linear in the size of the graphs.

Key words: Optimal graph matching. Error-correcting graph matching. Attributed relational graph. Graph edit distance. Subgraph isomorphism. Maximum common subgraph. Minimum common supergraph.

[1] Partially supported by the *Instituto de Cooperación Iberoamericana.*
[2] Partially supported by the Spanish DGES project PB96-0191-C02-02 and CICYT project TIC98-0949-C02-01 HEMOSS.

1 Introduction

Graphs are one of the most abstract and hence powerful structures for modeling problems and describing structured objects or concepts. The algorithms working on them have a wide variety of applications, from chemical studies to machine vision passing through information retrieval. These areas sometimes need to ask for the similarities of objects or represent a collection of similar objects using a single but representative one. In the latter case the maximum common subgraph can be used while in the former, a very popular way of measuring graphs similarities is used, the graph edit distance. This distance is based on a set of edit operations (deletion and insertion of both vertices and edges, and substitution of labels) with certain costs to take into account input data errors, and the problem of finding it is known as optimal error-correcting graph matching.

Recently, it has been shown that the maximum common subgraph of two graphs coincides with the optimal error-correcting graph matching between the two graphs, under certain assumptions on the underlying cost function [2]. As mentioned before, both problems have various applications in pattern recognition and computer vision, where objects are represented by attributed graphs and errors and distortions in the input graphs are taken care of. However, they do not provide an approximation allowing to combine several distorted input graphs due to missing structural information. Such an approximation could be provided by the minimum common supergraph.

In this article, the relationship between the maximum common subgraph and the minimum common supergraph of two graphs is established by means of simple constructions, which allow to obtain the maximum common subgraph from the minimum common supergraph, and vice versa. Thus, any algorithm for the computation of the maximum common subgraph [6,7,11] can be easily extended for the computation of the minimum common supergraph and hence under an appropriate cost function, any algorithm that implements error-correcting graph matching [3,5,8,9] can be used for the computation of both the maximum common subgraph and the minimum common supergraph. Furthermore, the maximum common subgraph and the minimum common supergraph are combined into a new graph distance metric, very close to graph edit distance but totally independent of edit costs.

2 Basic Definitions

Definition 1 *Let L be a finite alphabet of vertex and edge labels. A graph is a tuple $G = (V, E, \alpha)$, where*

- *V is the finite set of vertices,*
- *$E \subseteq V \times V$ is the finite set of edges,*
- *$\alpha : V \cup E \to L$ is the vertex and edge labeling function.*

Let $|G| = |V| + |E|$ denote the size of G. The empty graph such that $|G| = 0$ will be denoted by $\emptyset$.

If vertex and edge labels need to be explicitly distinguished, the set L can be partitioned into two disjoint sets or two different labeling functions could be used.

Attributed relational graphs, introduced in [10], extend the notion of labeled graph by allowing the value of several attributes as labels of vertices and edges. Since vertices and edges can be labeled in a uniform way, though, adding dummy attributes if needed, every vertex and every edge can be regarded as being labeled with the value of the same sequence of attributes. Therefore, attributed relational graphs are nothing but graphs labeled with tuples of values or with tuples of attribute-value pairs. In order to keep notations simple, and without loss of generality, only labeled graphs will be dealt with explicitly in this article.

Definition 2 *The union of two graphs $G_1 = (V_1, E_1, \alpha_1)$ and $G_2 = (V_2, E_2, \alpha_2)$, denoted by $G_1 \cup G_2$, is the graph $G = (V, E, \alpha)$ defined by*

- *$V = V_1 \cup V_2$,*
- *$E = E_1 \cup E_2$,*
- *$\alpha(x) = \begin{cases} \alpha_1(x) & \text{if } x \in V_1 \cup E_1, \\ \alpha_2(x) & \text{otherwise}, \end{cases}$*

if $\alpha_1(x) = \alpha_2(x)$ for all $x \in (V_1 \cap V_2) \cup (E_1 \cap E_2)$.

Definition 3 *A graph $G_1 = (V_1, E_1, \alpha_1)$ is a subgraph of a graph $G_2 = (V_2, E_2, \alpha_2)$, denoted by $G_1 \subseteq G_2$, if*

- *$V_1 \subseteq V_2$,*
- *$E_1 \subseteq E_2$,*
- *$\alpha_1(x) = \alpha_2(x)$ for all $x \in V_1 \cup E_1$.*

A graph G_2 is called a supergraph of G_1 if $G_1 \subseteq G_2$.

Definition 4 *Let $G = (V, E, \alpha)$ be a graph, V_2 a set of vertices, and $f : V_1 \to V_2$ a bijective function, where $V_1 \subseteq V$ and $V_1 \cap V_2 = \emptyset$. A renaming of G by f, denoted by $f(G)$, is the graph $G_f = (V_f, E_f, \alpha_f)$ defined by*

- *$V_f = V_2 \cup (V \setminus V_1)$,*

- $E_f = \{(f'(u), f'(v)) \mid (u,v) \in E\}$,
- $\alpha_f(f'(v)) = \alpha(v)$ *for all* $v \in V$,
- $\alpha_f(e) = \alpha(e)$ *for all* $e \in E$.

where $f' : V \to V_f$ *is defined by* $f'(v) = \begin{cases} f(v) & \text{if } v \in V_1, \\ v & \text{otherwise.} \end{cases}$

Definition 5 *Two graphs G_1 and G_2 are isomorphic, denoted by $G_1 \cong G_2$, if there is a renaming f of G_1 such that $f(G_1) = G_2$, and in this case it is said that $f : G_1 \to G_2$ is a graph isomorphism.*

Definition 6 *A graph $\hat{G}$ is a common subgraph of two graphs G_1 and G_2 if there exist subgraphs $\hat{G}_1 \subseteq G_1$ and $\hat{G}_2 \subseteq G_2$ such that $\hat{G} \cong \hat{G}_1 \cong \hat{G}_2$. It is maximum if there is no other common subgraph of G_1 and G_2 larger than $\hat{G}$.*

Definition 7 *A graph $\check{G}$ is a common supergraph of two graphs G_1 and G_2 if there exist graphs $\check{G}_1 \subseteq \check{G}$ and $\check{G}_2 \subseteq \check{G}$ such that $\check{G}_1 \cong G_1$ and $\check{G}_2 \cong G_2$. It is minimum if there is no other common supergraph of G_1 and G_2 smaller than $\check{G}$.*

It is worth note that according to the above definitions, a maximum common subgraph (minimum common supergraph) contains the largest (smallest) possible number of both, vertices and edges.

3 A Different Graph Difference

The constructions relating the maximum common subgraph and the minimum common supergraph are based on a particular notion of graph difference, with the purpose of eliminating the common structural information of two graphs. That is why this difference is defined using the difference of the sets of edges rather than vertices as usual, and it needs to operate over *full graphs* (graphs having no isolated vertices) in order to provide a complete, proper, and simple set of properties.

Definition 8 *A graph $G = (V, E, \alpha)$ is full if $\{u, v \mid (u,v) \in E\} = V$.*

Definition 9 *Let G_1 and G_2 be two graphs with $G_1 \subseteq G_2$. The complementary graph difference of G_1 and G_2, denoted by $G_2 \ominus G_1$, is the graph $G = (V, E, \alpha)$ given by*

- $E = E_2 \setminus E_1$,
- $V = V_2 \setminus V_1 \cup \{u, v \mid (u,v) \in E_2 \setminus E_1\}$,
- $\alpha(x) = \alpha_2(x)$ *for all* $x \in V \cup E$.

The operation of complementary graph difference is well defined, and using this definition, the following basic properties can be easily checked.

Proposition 10 *Let G_1, G_2 and G be graphs such that $G_1 \subseteq G_2$. Then*

(1) $G_2 = G_1 \cup (G_2 \ominus G_1)$,
(2) $(G_2 \ominus (G_2 \ominus G_1)) \subseteq G_1$, *and both have the same set of edges,*
(3) $G_2 \ominus G_1 = \emptyset$ *if and only if* $G_2 = G_1$,
(4) $|G_2 \ominus G_1| = |G_2| - |G_1| + |G_1 \cap (G_2 \ominus G_1)|$.

Example 11 *The complementary graph difference $G_2 \ominus G_1$ of a graph G_2 and a subgraph $G_1 \subseteq G_2$ is illustrated by the following figure:*

$G_1 \subseteq G_2$

$G_2 \ominus G_1$

where $V_1 = \{2,3,4\}$ *and* $V_2 = \{1,2,3,4\}$ *but* $G_2 \ominus G_1$ *has* $\{1,2,3\}$ *as set of vertices, instead of just* $\{1\} = V_2 \setminus V_1$. *Moreover,* $G_2 = G_1 \cup (G_2 \ominus G_1)$, $(G_2 \ominus (G_2 \ominus G_1)) = G_1$, *and* $G_1 \cap (G_2 \ominus G_1) \neq \emptyset$.

Also under this notion of graph difference, the following properties of graph union with respect to complementary graph difference hold.

Proposition 12 *Let G_1, G_2 and G_3 be full graphs with $E_1 \cap E_2 = \emptyset$ and $G_3 \subseteq G_2$. Then $(G_1 \cup G_2) \ominus G_3 = G_1 \cup (G_2 \ominus G_3)$.*

Proposition 13 *Let G_1, G_2 and G_3 be full graphs with $G_1 \subseteq G_2 \subseteq G_3$. Then $G_3 \ominus (G_2 \ominus G_1) = (G_3 \ominus G_2) \cup G_1$.*

Note that in the case of Proposition 12, when $E_1 \cap E_2 = \emptyset$, $(E_1 \cup E_2) \setminus E_3 = E_1 \cup (E_2 \setminus E_3)$ by definition of set difference. On the other hand, using the definition of complementary graph difference to replace G_3 by $(G_3 \ominus G_2) \cup (G_2 \ominus G_1) \cup G_1$, the required results in Proposition 13 can be easily obtained.

It is also interesting to see how with additional restrictive conditions over isolated vertices, these two propositions also hold for general graphs and in particular, Proposition 12 holds when the isolated vertices of G_3 are not isolated in G_1 and in G_2.

Proposition 14 *Let G_1, G_2 and G_3 be graphs with $E_1 \cap E_2 = \emptyset$, $G_3 \subseteq G_2$ and $V_1 \cap V_3 \subseteq \{u, v | (u,v) \in E_1 \cup (E_2 \ominus E_3)\}$. Then $(G_1 \cup G_2) \ominus G_3 = G_1 \cup (G_2 \ominus G_3)$.*

4 From Maximum Common Subgraph to Minimum Common Supergraph and Back

Using the above properties of complementary graph difference and given the maximum common subgraph of two graphs, the minimum common supergraph can be obtained by means of a simple construction.

Theorem 15 *Let $\hat{G}$ be a maximum common subgraph of two graphs G_1 and G_2, $\hat{f}_1 : V_1 \to \hat{V}$ a graph isomorphism between $\hat{G}$ and $\hat{G}_1$, and $\hat{f}_2 : V_2 \to \hat{V}$ a graph isomorphism between $\hat{G}$ and $\hat{G}_2$. Then, a minimum common supergraph $\check{G}$ of G_1 and G_2 is given by $\check{G} \cong G_1 \cup ((\hat{f}_2^{-1} \circ \hat{f}_1)(G_2) \ominus (\hat{f}_2^{-1} \circ \hat{f}_1)(\hat{G}_2))$.*

Proof Let $G = G_1 \cup ((\hat{f}_2^{-1} \circ \hat{f}_1)(G_2) \ominus (\hat{f}_2^{-1} \circ \hat{f}_1)(\hat{G}_2))$, and let $\check{G}$ be a minimum common supergraph of G_1 and G_2. It is not difficult to see that G is a common supergraph of G_1 and G_2, and

$$\begin{aligned} |G| &= |G_1| + |G_2| - |\hat{G}| - |\check{G}_1 \cap (\check{G}_2 \ominus \hat{G})| + |\hat{G} \cap (\check{G}_2 \ominus \hat{G})| \\ &= |G_1| + |G_2| - |\hat{G}| \,. \end{aligned}$$

Consequently, the inequalities $|G| \geqslant |\check{G}|$, $|G_1| + |G_2| - |\hat{G}| \geqslant |\check{G}|$ and $|G_1| + |G_2| - |\check{G}| \geqslant |\hat{G}|$ hold. But, since $\check{G}_1 \cup \check{G}_2 = \check{G}$, $\hat{G}' = \check{G}_1 \ominus (\check{G} \ominus \check{G}_2)$ is a common subgraph of $\check{G}_1$ and $\check{G}_2$, and the inequality $|\hat{G}'| = |G_1| + |G_2| - |\check{G}| \leqslant |\hat{G}|$ also holds. Therefore, $|G_1| + |G_2| - |\check{G}| = |\hat{G}|$, $|G| = |\check{G}|$, and then G is a minimum common supergraph of G_1 and G_2. $\square$

Notice that in the previous theorem, any of the expressions

$$\begin{aligned} \check{G} &\cong G_1 \cup ((\hat{f}_2^{-1} \circ \hat{f}_1)(G_2) \ominus (\hat{f}_2^{-1} \circ \hat{f}_1)(\hat{G}_2)) \\ &\cong G_2 \cup ((\hat{f}_1^{-1} \circ \hat{f}_2)(G_1) \ominus (\hat{f}_1^{-1} \circ \hat{f}_2)(\hat{G}_1)) \\ &\cong \hat{G} \cup (\hat{f}_1(G_1) \ominus \hat{f}_1(\hat{G}_1)) \cup (\hat{f}_2(G_2) \ominus \hat{f}_2(\hat{G}_2)) \end{aligned}$$

can be used.

Now, given a minimum common supergraph of two graphs, a common subgraph with the same number of edges as a maximum common subgraph can also be obtained using the following result.

Theorem 16 *Let $\check{G}$ be a minimum common supergraph of two graphs G_1 and G_2. Then $\check{G}_1 \ominus (\check{G} \ominus \check{G}_2)$ is a common subgraph of G_1 and G_2 with a maximum number of edges.*

Proof As a matter of fact, by Theorem 15 it holds that $\check{G}_1 \ominus (\check{G} \ominus \check{G}_2) = \check{G}_1 \ominus ((\check{G}_2 \cup (\check{G}_1 \ominus \hat{G})) \ominus G_2)$, where $\hat{G}$ is a maximum common subgraph of $\check{G}_1$ and $\check{G}_2$. Further, since the graphs $\check{G}_2$ and $(\check{G}_1 \ominus \hat{G})$ have no common edges and the expression $\check{G}_2 \cap (\check{G}_1 \ominus \hat{G}) = \hat{G} \cap (\check{G}_1 \ominus \hat{G}) \subseteq \{u, v | (u, v) \in \check{E}_1\}$ holds, it follows by Proposition 12 that $\check{G}_1 \ominus ((\check{G}_2 \cup (\check{G}_1 \ominus \hat{G})) \ominus \hat{G}_2) = \check{G}_1 \ominus ((\check{G}_1 \ominus \hat{G}) \cup (\check{G}_2 \ominus \check{G}_2))$ and finally, $\check{G}_1 \ominus (\check{G}_1 \ominus \hat{G}) = G \subseteq \hat{G}$ with the required condition of having the same number of edges, by Proposition 10. $\square$

Again, notice that in the previous theorem, the equivalent expression $\check{G}_2 \ominus (\check{G} \ominus \check{G}_1)$ can be used instead. Notice also that the previous construction does not give a maximum common subgraph, because $\check{G}_1 \ominus (\check{G} \ominus \check{G}_2)$ is maximal in the number of edges but not in the number of vertices.

5 A New Graph Distance Metric

The concrete edit costs for a graph edit distance are strong application-dependent and hence there is no general and effective method to obtain them, except by hand. So, instead of relying on fundamental graph edit operations, a graph distance metric has been proposed in [1,4] which is based on the maximum common subgraph of two graphs G_1 and G_2, namely $\delta(G_1, G_2) = 1 - |\hat{G}| / \max(|G_1|, |G_2|)$. Being a metric, a distance allows to establish a partial order over the objects of interest, and its properties are a useful guide for the searching task. That is why in many applications it is desirable to have a metric for measuring object similarities.

This section introduces a new graph distance metric that is based on both, maximum common subgraph and minimum common supergraph, and therefore takes into consideration not only superfluous but also missing structural information in the input graphs. Moreover, it will be shown that this new metric has a very close relationship to those distances defined using the notion of error-correcting transformation.

Theorem 17 *Let $\hat{G}$ and $\check{G}$ be respectively a maximum common subgraph and a minimum common supergraph of G_1 and G_2. The graph distance defined by $\delta(G_1, G_2) = |\check{G}| - |\hat{G}|$ is a metric.*

Proof It has to be shown that

(1) $\delta(G_1, G_2) \geqslant 0$, and $\delta(G_1, G_2) = 0$ if and only if $G_1 \cong G_2$;
(2) $\delta(G_1, G_2) = \delta(G_2, G_1)$; and
(3) $\delta(G_1, G_3) \leqslant \delta(G_1, G_2) + \delta(G_2, G_3)$

for any graphs G_1, G_2, and G_3. But notice that $\delta(G_1, G_2) = |G_1| + |G_2| - 2|\hat{G}_{12}|$, so the two first conditions have a straightforward proof and the third one, $\delta(G_1, G_3) \leqslant \delta(G_1, G_2) + \delta(G_2, G_3)$, is equivalent to proving the inequality $|G_2| + |\hat{G}_{13}| - |\hat{G}_{12}| - |\hat{G}_{23}| \geqslant 0$.

Let $\hat{G}$ and $\check{G}$ be respectively a maximum common subgraph and a minimum common supergraph of $\hat{G}_{12}$ and $\hat{G}_{23}$. Then,

$$\begin{aligned} |G_2| + |\hat{G}_{13}| - |\hat{G}_{12}| - |\hat{G}_{23}| &= |G_2| + |\hat{G}_{13}| - (|\hat{G}_{12}| + |\hat{G}_{23}| - |\hat{G}|) - |\hat{G}| \\ &= |G_2| - |\check{G}| + |\hat{G}_{13}| - |\hat{G}| \,, \end{aligned}$$

and this is a non-negative expression, by the fact that $\check{G} \subseteq G_2$ and $\hat{G} \subseteq \hat{G}_{13}$. $\square$

Definition 18 *Let $G_1 = (V_1, E_1, \alpha_1)$ and $G_2 = (V_2, E_2, \alpha_2)$ be two graphs. An error-correcting graph matching from G_1 to G_2 is a pair $\zeta = \langle G, f_G \rangle$, where $G = (V, E, \alpha) \subseteq G_1$, (V, E, α_c) is a common subgraph of (V_1, E_1, α_c) and (V_2, E_2, α_c), $f_G : V \to \hat{V}_2$ is the corresponding isomorphism, and α_c is a constant function.*

The pair $\zeta = \langle G, f_G \rangle$ encodes the edit operations of insertion, deletion and substitution of both vertices and edges. So, when a graph transformation is performed over G_1 and G_2 using the error-correcting graph matching ζ, any vertex $u \in V$ (edge $e \in E$) will be substituted by the vertex $f_G(v) \in \hat{V}_2$ (edge $f_G(e) \in \hat{E}_2$), replacing the vertex (edge) label if necessary. (This implies that identical substitutions will always have zero cost.) Furthermore, any edge of $E_1 \setminus E$ will be deleted followed by the deletion of the vertices of $V_1 \setminus V$, and the same with the insertion of any vertex of $V_2 \setminus \hat{V}_2$ and edge of $E_2 \setminus \hat{E}_2$.

Definition 19 *A cost function C is a vector of non-negative real functions $(c_{\text{nd}}(v), c_{\text{ni}}(v), c_{\text{ns}}(v_1, v_2), c_{\text{ed}}(e), c_{\text{ei}}(e), c_{\text{es}}(e_1, e_2))$, where each component represents a cost for the vertex insertion, vertex deletion, vertex substitution, edge insertion, edge deletion and edge substitution, respectively, on graphs. The cost of an error-correcting graph matching $\zeta = \langle G, f_G \rangle$ from G_1 to G_2 under C is*

$$\begin{aligned} C(\zeta) = &\sum_{v \in V_1 \setminus V} c_{\text{nd}}(v) + \sum_{v \in V_2 \setminus V} c_{\text{ni}}(v) + \sum_{v \in V} c_{\text{ns}}(v, f_G(v)) + \\ &\sum_{e \in E_1 \setminus E} c_{\text{ed}}(e) + \sum_{e \in E_2 \setminus E} c_{\text{ei}}(e) + \sum_{e \in E} c_{\text{es}}(e, f_G(e)) \,. \end{aligned}$$

In other words, the cost of an error-correcting graph matching ζ is the sum of the costs of all edit operations implied by ζ.

Definition 20 *The edit distance between two graphs G_1 and G_2 is defined by $dist(G_1, G_2) = \min\{C(\zeta) \mid \zeta$ error-correcting graph matching from G_1 to $G_2\}$. Any error-correcting graph matching ζ from G_1 to G_2 such that $dist(G_1, G_2) = C(\zeta)$ will be called an optimal error-correcting graph matching.*

This edit distance is a slight modification of the one defined in [2]. Hence, the following propositions hold using almost the same proofs given in that article.

Proposition 21 *Let G_1 and G_2 be graphs, $G \subseteq G_1$ a maximum common subgraph of G_1 and G_2, f an isomorphism between G and $\hat{G}_2$ and $C = (c_{\rm nd}, c_{\rm ni}, c_{\rm ns}, c_{\rm ed}, c_{\rm ei}, c_{\rm es})$ a cost function, where $c_{\rm nd}, c_{\rm ni}, c_{\rm ed}, c_{\rm ei}$ are constant functions such that for all $v_1 \in V_1$ and $v_2 \in V_2$ with $\alpha_1(v_1) \neq \alpha_2(v_2)$, $c_{\rm nd}(v_1) + c_{\rm ni}(v_2) \leqslant c_{\rm ns}(v_1, v_2)$, and for all $e_1 \in E_1$ and $e_2 \in E_2$ with $\alpha_1(e_1) \neq \alpha_2(e_2)$, $c_{\rm ed}(e_1) + c_{\rm ei}(e_2) \leqslant c_{\rm es}(e_1, e_2)$. Then, $\langle G, f\rangle$ is an optimal error-correcting graph matching from G_1 to G_2.*

Proposition 22 *Let G_1 and G_2 be graphs and let $C = (c_{\rm nd}, c_{\rm ni}, c_{\rm ns}, c_{\rm ed}, c_{\rm ei}, c_{\rm es})$ be a cost function where $c_{\rm nd}, c_{\rm ni}, c_{\rm ed}, c_{\rm ei}$ are constant functions such that for all $v_1 \in V_1$ and $v_2 \in V_2$ with $\alpha_1(v_1) \neq \alpha_2(v_2)$, $c_{\rm nd}(v_1) + c_{\rm ni}(v_2) < c_{\rm ns}(v_1, v_2)$, and for all $e_1 \in E_1$ and $e_2 \in E_2$ with $\alpha_1(e_1) \neq \alpha_2(e_2)$, $c_{\rm ed}(e_1) + c_{\rm ei}(e_2) < c_{\rm es}(e_1, e_2)$. Then, for any optimal error-correcting graph matching $\langle G, f_G\rangle$ from G_1 to G_2, G is a maximum common subgraph of G_1 and G_2.*

By all these results, the following theorem has a direct proof.

Theorem 23 *Let $C = (c, c, c_{\rm ns}, c, c, c_{\rm es})$ be a cost function, where c is a constant function and for all $v_1 \in V_1$ and $v_2 \in V_2$ with $\alpha_1(v_1) \neq \alpha_2(v_2)$, it holds that $c_{\rm ns}(v_1, v_2) > 2c$ and $c_{\rm es}(e_1, e_2) > 2c$. Then, for any graphs G_1 and G_2, $dist(G_1, G_2) = c\delta(G_1, G_2)$.*

6 Conclusion

In this article, the relationship between maximum common subgraphs and minimum common supergraphs of two graphs has been established by means of basic constructions and a new graph metric was defined, very close to a graph edit distance.

The simplicity of these constructions shows that problems involving the search for maximum commonality and minimum combination of structural information on attributed relational graphs are computationally equivalent, and allows to easily extend any algorithm for the computation of the maximum common subgraph, and also any algorithm that implements error-correcting graph matching (when run under an appropriate cost function) to the computation of the minimum common supergraph, and the extension only takes time linear in the size of the graphs.

All these formulations lie on a particular graph difference that operates over full graphs (graphs having no isolated vertices). However, this restriction is not a serious one. In fact, the properties used for proving the main results of this article hold for general attributed relational graphs, and therefore an implementation of an algorithm for the computation of a maximum common subgraph of two graphs from a minimum common supergraph can be easily and efficiently obtained.

References

[1] H. Bunke. On a relation between graph edit distance and maximum common subgraph. *Pattern Recogn. Lett.*, 18(8):689–694, 1997.

[2] H. Bunke. Error-correcting graph matching: On the influence of the underlying cost function. *IEEE T. Pattern Anal.*, 21(9):917–922, 1999.

[3] H. Bunke and G. Allermann. Inexact graph matching for structural pattern recognition. *Pattern Recogn. Lett.*, 1(4):245–253, 1983.

[4] H. Bunke and K. Schearer. A graph distance metric based on the maximal common subgraph. *Pattern Recogn. Lett.*, 19:255–259, 1998.

[5] S. Gold and A. Rangarajan. A graduated assignment algorithm for graph matching. *IEEE T. Pattern Anal.*, 18(4):377–388, 1996.

[6] G. Levi. A note on the derivation of maximal common subgraphs of two directed or undirected graphs. *Calcolo*, 9:1–12, 1972.

[7] J. J. McGregor. Backtrack search algorithms and the maximal common subgraph problem. *Software—Practice and Experience*, 12:23–34, 1982.

[8] B. T. Messmer and H. Bunke. A new algorithm for error-tolerant subgraph isomorphism detection. *IEEE T. Pattern Anal.*, 20(5):493–504, 1998.

[9] L. G. Shapiro and R. M. Haralick. Structural descriptions and inexact matching. *IEEE T. Pattern Anal.*, 3(5):504–519, 1981.

[10] W. H. Tsai and K.-S. Fu. Error-correcting isomorphism of attributed relational graphs for pattern analysis. *IEEE T. Syst. Man Cyb.*, 9(12):757–769, 1979.

[11] A. K. C. Wong and F. A. Akinniyi. An Algorithm for the Largest Common Subgraph Isomorphism using the Implicit Net. In *Proc. Int. Conf. Systems, Man and Cybernetics*, volume 1, pages 197–201. IEEE, 1983.