

An Heuristic for Graph Symmetry Detection

Hubert de Fraysseix

CNRS UMR 8557

E.H.E.S.S.

54 Bd Raspail

75006 Paris

France

hf@ehess.fr

Abstract. We give a short introduction to an heuristic to find automorphisms in a graph such as axial, central or rotational symmetries. Using techniques of factorial analysis, we embed the graph in an Euclidean space and try to detect and interpret the geometric symmetries of the embedded graph. It has been particularly developed to detect axial symmetries.

1 Introduction

Testing whether a graph has any axial (rotational, central, respectively) symmetry is a NP-complete problem [15]. Some restrictions (central symmetry with exactly one fixed vertex and no fixed edge) are polynomially equivalent to the graph isomorphism test. Notice that this latter problem is not known to be either polynomial or NP-complete in general. For isomorphisms, several heuristics are known (e.g. [4]) and several restrictions leads to efficient algorithms: linear time isomorphism test for planar graphs [11] and interval graphs [14], polynomial time isomorphism test for fixed genus [18, 7], k -contractible graphs [20] and pairwise k -separable graphs [19], linear axial symmetry detection for planar graphs with a given embedding. Symmetry detection and display for trees and, more generally, for outerplanar graphs, serie-parallel digraphs and planar graphs may be found in [16][17][9][10].

In the context of automatic generation of industrial diagrams, the aim of symmetry display is not only to get a 'nice' drawing but also to highlight some of the semantic of the design. Usually the given networks are not planar and even so, due to other constraints (even aesthetics ones), crossings are very common.

We present here the main outlines of an heuristic to detect symmetries. A mathematical justification is to appear in [5], where we propose an exact algorithm. A specific version will soon be integrated in PICTEL, our diagram drawing module of CATIA®¹ but has already been integrated in our research graph drawing toolkit PIGALE².

¹ CATIA is a registered trade mark of IBM

² PIGALE is a Public Implementation of a Graph Algorithm Library and Editor which is available by anonymous FTP at <ftp://pr.cams.ehess.fr/pub/pigale.tgz>

Our main idea to exhibit the symmetries of a graph (planar or not) is to define a distance among the vertices of the graph and embed it isometrically in \mathbb{R}^{n-1} , where n is the order of the graph. Then detecting the symmetries of the graph mainly reduces to finding the geometric symmetries of the embedded graph and checking that they do corresponds to symmetries of the given graph.

2 Automorphisms and Symmetries

An *automorphism* of $G = (V, E)$ is a permutation σ of V and E , such that:

- $\sigma(E) = E$,
- $\sigma(V) = V$,
- $e \in E$ is incident to $v \in V \iff \sigma(e)$ is incident to $\sigma(v)$

The usual meaning given to axial, central and rotational symmetry is strongly related to a drawing of a graph. We shall say that such symmetries exist if and only if there exists a drawing of the graph in the plane (where all vertices are represented by different points and edges by lines which may cross) in such a way that a usual geometric axial, central or rotational symmetry appears. A *central symmetry* is an involutive automorphism $\sigma \in \text{Aut}(G)$, such that the fixed points set of σ is either included in E or reduced to a single vertex.

A *rotational symmetry* is an automorphism $\sigma \in \text{Aut}(G)$, which cycles are all of the same length (greater than two) and such that the fixed points set of σ is either empty or reduce to a single vertex. The length of the cycles is the *order* of the symmetry and the rotational symmetry will be said to be *even* or *odd* depending on the evenness of its order.

An *axial symmetry* is an involutive automorphism $\sigma \in \text{Aut}(G)$, such that the subgraph of G induced by the fixed points set of σ is embeddable on a line.

A *planar symmetry* is a symmetry which appears as a geometric symmetry for some planar embedding of the graph.

Remark 1. An even rotational symmetry define axial symmetries but an odd one may define none.

3 The Embedding Model

We shall recall some very basic methods used in *Factorial Analysis* applied to the study of statistical data [3][12].

3.1 Euclidean Semi-Distances

Let X be a set of n points. A *semi-distance* d on X is a mapping which satisfies all the axioms of a distance but one: two distinct points are allowed to be at null semi-distance. The semi-distance d is *Euclidean*, if there exists an isometry from X to \mathbb{R}^{n-1} : the restriction of the usual distance of \mathbb{R}^{n-1} to the image of X coincides with the given semi-distance d on X .

Given a semi-distance d on a set X , it is classical to define the corresponding inner-product matrix W define by the formula:

$$W_{i,j} = \frac{1}{2}(d^2(v_i, \cdot) + d^2(v_j, \cdot) - d^2(\cdot, \cdot)) \tag{1}$$

where

$$d^2(v_i, \cdot) = \frac{1}{n} \sum_{k=1}^n d^2(v_i, v_k)$$

$$d^2(\cdot, \cdot) = \frac{1}{n} \sum_{i=1}^n d^2(v_i, \cdot)$$

It is well known that the semi-distance d is Euclidean if and only if the matrix W defines a positive semi-definite bilinear form [1][8][13], that is all its eigenvalues are positive or null. The rank of W (i.e. the maximum of non null eigenvalues) is bounded by $n - 1$, the maximal dimension of the vector space generated by n points.

If we denote by F_1, \dots, F_p an orthonormal basis of eigenvectors associated to the strictly positive eigenvalues $\lambda_1, \dots, \lambda_p$, we have:

$$d^2(v_i, v_j) = \sum_{k=1}^p \lambda_k (F_{i,k} - F_{j,k})^2 \tag{2}$$

3.2 Defining a Semi-Distance in a Graph

Seeking a *natural* semi-distance between vertices of a graph, one could consider the length of the shortest paths joining them. But even the complete graph on 4 vertices with one edge deleted proves that shortest paths do not define an Euclidean semi-distance.

Many mathematical statisticians have define Euclidean semi-distances on abstract sets. But from previous work, while supervising P. Kunz’s thesis [13] on the problem of partitioning large graphs into a given number of subgraphs [6], the so called Czekanovski-Dice semi-distance seemed to be very appropriate to reveal the structure of a graph. With respect to that semi-distance, two vertices are close to each other if they have many common neighbors: For each pair (v_i, v_j) of vertices, we denote N_i the set of the neighbors of v_i :

$$N_i = \{v_k \in V, (v_i, v_j) \in E\} \cup \{v_i\}$$

Then, the semi-distance d is defined by:

$$d^2(v_i, v_j) = 1 - 2 \cdot \frac{|N_i \cap N_j|}{|N_i| + |N_j|} \tag{3}$$

Remark that two adjacent vertices having the same neighbors are at semi-distance 0 and that two non adjacent vertices having no common neighbor are at semi-distance 1. Notice that the semi-distances are obviously preserved by any automorphism of the graph.

3.3 Embedding and Projecting the Graph on Planes

The first step of the heuristic is to compute the Czekanovski-Dice semi-distance among the vertices of the graph and the corresponding inner-product matrix W . As the Czekanovski-Dice semi-distance is Euclidean, we can diagonalise W whose eigenvalues are all positive or null. Unless G is a complete graph, the rank of W is not null, and if G has more than 2 vertices its rank is at least 2, but if there are many pairs of vertices with the same neighbors, the rank of W decreases and we get many null eigenvalues.

We then compute an orthonormal basis and compute the p coordinates of all the vertices corresponding to the p strictly positive eigenvalues. We then sort the axis in decreasing order according to the value of their corresponding eigenvalue. Using the terminology of mechanics, the origin is the inertia center, the axis are the principal axis of the set of points and the inertia along each axis decreases with respect to that order. The main tool of Factorial Analysis is precisely to study the projections of the points on the planes defined by the first axis which reflects the main semi-distances structure.

So the next step of the heuristic then consists in looking at the geometric symmetries that may appear when projecting the points on the planes defined by the first axis.

4 The Heuristic by Examples

Paradoxically, it is far easier to find a symmetry in a graph with a small automorphism group (when one exists!). If it is not the case, there are many geometric symmetries in the embedded graph and many equal eigenvalues (e.g the Petersen graphs has, among its nine eigenvalues, only two distinct ones). If an eigenvalue has a multiplicity greater than one, there is no canonical way to compute the corresponding eigenvectors and heuristics have to be designed to select the most appropriate ones such that in some plane defined by them a geometrical symmetry appears.

4.1 All the Eigenvalues are Distinct

This is the easiest case, as all the principal axis are uniquely determined. In all the cases we encountered, it is only necessary to check for a geometric axial symmetry of the graph projected on the first two principal axis (x, y) , or to check for the identity. Depending on the number of edges between the two isomorphic graphs defined by the axial symmetry, the geometric symmetry appears on the y -axis or the x -axis. The points with null coordinates are the fixed points, and the other points split into two sets (the one with positive x -coordinates and the ones with x -negative coordinates). Usually the geometric symmetry is defined by the y axis: the fixed points are the points with null x -coordinate, and the two isomorphic graphs are defined by the points of positive (resp. negative x -coordinates). If no two points have the same x -coordinate, the identification of

the corresponding symmetric vertices is trivial. Otherwise the third axis usually breaks the ambiguities. In a few cases, the edges joining the fixed points, are not a chain subgraph (i.e. the axial symmetry detection fails) but we do hope to solve this problem by doing a closer analysis of the fixed points set. We never found an example where geometric symmetry did not correspond to an automorphism of the graph. Figure 1 is a typical example where the first eigenvalues are distinct. The image on the left side represents the projection on the first two principal axis, and on the right side we show the graph where the vertices have been relabelled according to the symmetry. The drawings of the planar graphs use the so-called Tutte-Circle algorithm of PIGALE (the postscript files were created by the LEDA library).

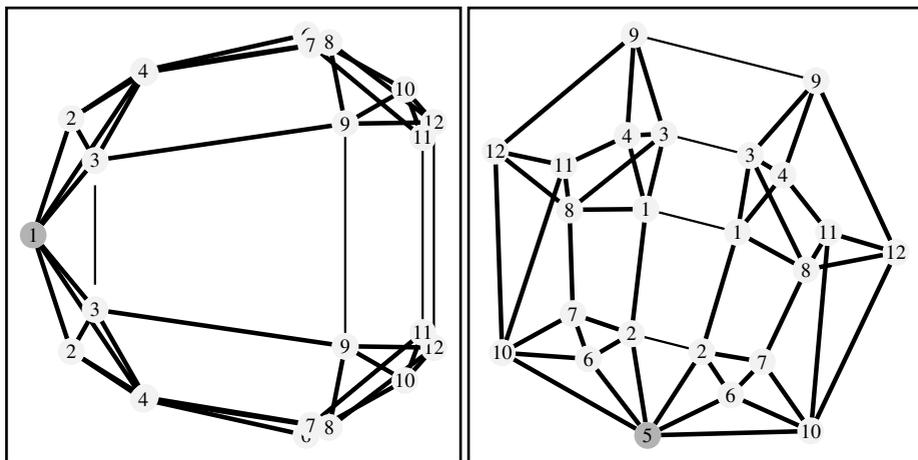


Fig. 1. A non-planar graph with only an axial symmetry

4.2 The First Eigenvalue has Multiplicity Two

In such a case, we highly suspect the existence of a central symmetry or a rotational symmetry. For a central symmetry, we only have to check that there exists at most one point with x and y null coordinates and that the points have a geometric central symmetry.

Then we try to exhibit a rotational symmetry. This is also very simple: we consider the points which are at maximal distance of the origin and sort them according to their polar angle and we try to match them by rotations (matching-pattern type algorithm). If we succeed, we check that the matching extends to all the points.

If we find a central or rotational symmetry, we try to exhibit an axial symmetry. As the two first eigenvalues are equal, the position of the first axis depend on the labelling of the vertices and has nothing to do with its structure. So we

really need an heuristic to find out a rotation of the first two axis such that a geometrical axial symmetry appears along one of those axis in the plane they define. If we do have an even rotational symmetry, the problem is extremely simple. Otherwise it seems rather difficult and we are far from having an algorithm that we could guaranty. Nevertheless, we do have an heuristic that has been efficient in all the cases we met. It is probably due to the fact that we only consider rotational symmetries when the first two eigenvalues are equal.

Figure 2, Figure 3, Figure 4 gives examples where the first eigenvalue is multiple. The Factorial Analysis drawings (left of the pictures) are done after the correct rotation of the axis has been applied. (We only show one or two symmetries found by the heuristic).

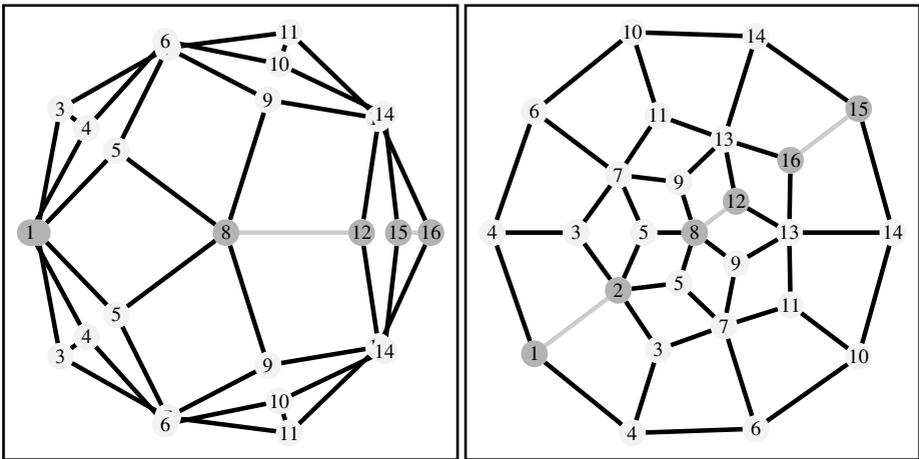


Fig. 2. A planar graph with rotational and axial symmetries

4.3 General Case

If the first eigenvalue has multiplicity one, and the second one has multiplicity two, we are more or less in the previous situation.

If all the eigenvalues of a graph are equal, we get no information from any projection, but we suspect that the graph is vertex transitive (i.e. for any pair of vertices there is an automorphism matching them).

As we said earlier, if two adjacent vertices have the same neighbors, they have the same coordinates and so the null eigenvalue has a multiplicity greater than one. If the null eigenvalue has a big multiplicity, we probably could find interesting automorphisms.

Otherwise, we do have examples where our heuristic completely fails (e.g. Petersen graph, complete bipartite graphs etc).

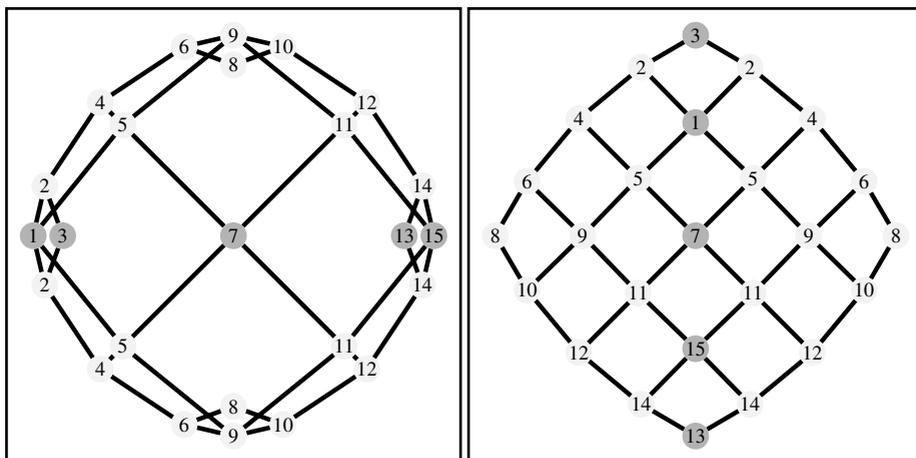


Fig. 3. A grid graph with one of its axial symmetries

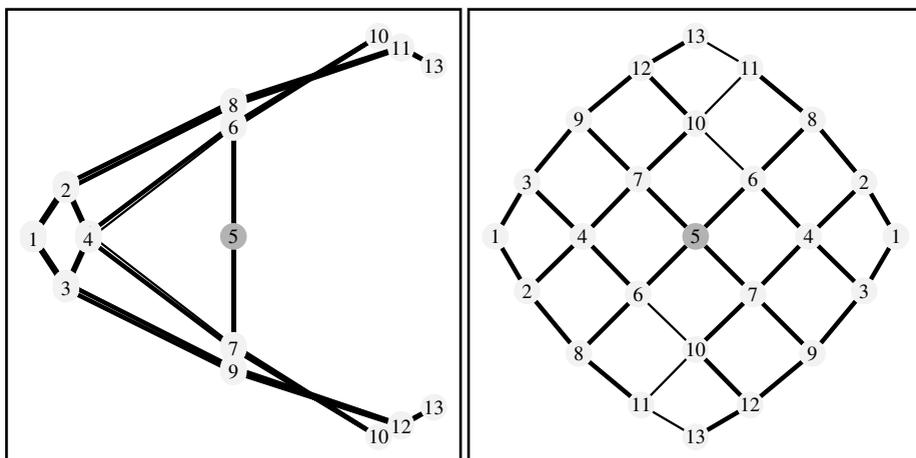


Fig. 4. A grid graph with a central symmetry

4.4 More Pictures

The following pictures, represent famous graphs taken from *Examples and Counterexamples in Graph Theory* [2]:

- Figure 5 represents the Folkman graph.
- Figure 6 On the left, a graph whose automorphism group is the Klein 4-group, and a graph whose automorphism group is C_6 .

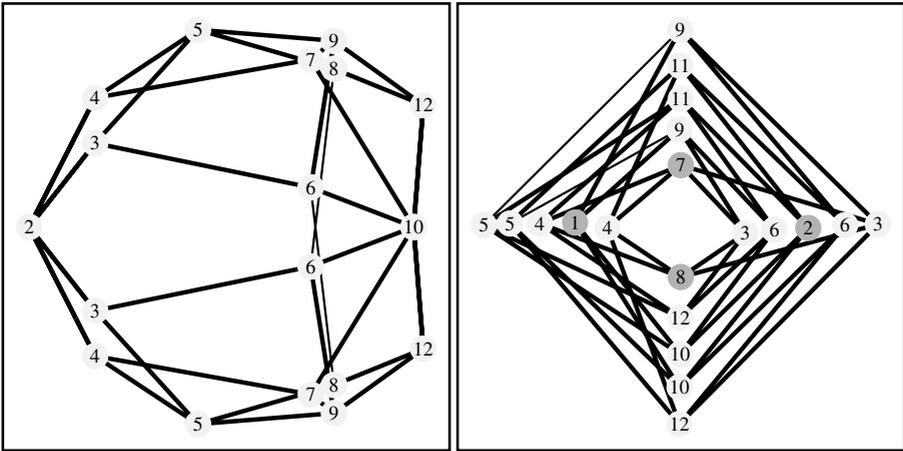


Fig. 5. The Folkman graph and one its Factorial Analysis projection

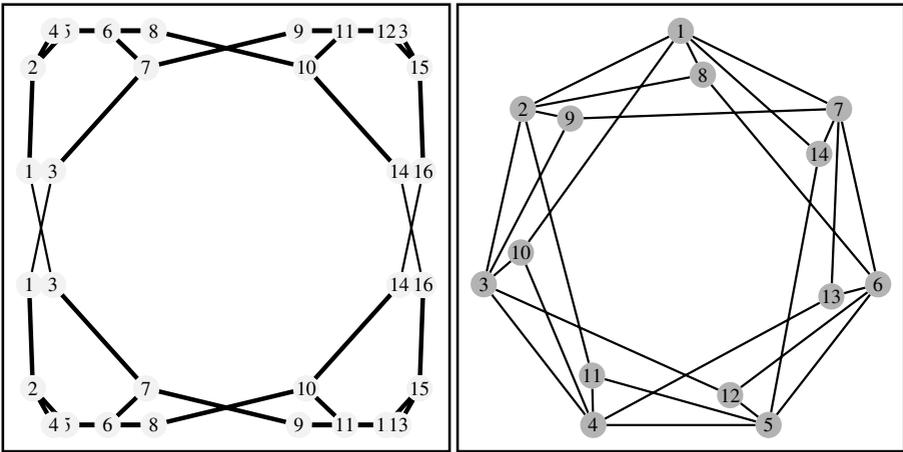


Fig. 6. Two other Factorial Analysis projections from [2]

4.5 Open Questions

- Can two non-isomorphic graphs have the same semi-distances among their vertices?
- Is there a graph with a rotational or central symmetry that does not have at least two equal eigenvalues?
- Is there a graph having all its eigenvalues equal which is not vertex transitive (i.e. for any pair of vertices there is an automorphism matching them)?
- If a graph have two adjacent vertices with the same neighbors, the null eigenvalue has a multiplicity greater than one. Characterize the other configurations that imply that the null eigenvalue has a multiplicity greater than one.

4.6 Concluding Remarks

Sometimes, a geometric axial symmetry only reveals an automorphism as the set of fixed points may not be a subgraph of a chain. This happens very often when the graph is a tree, but we shall implement the linear time algorithm to exhibit axial symmetries in such cases. As stated before, the heuristic often fails if the first eigenvalues have a great multiplicity. In such cases a higher complexity algorithm should be devised. Otherwise this heuristic seems very effective and do find symmetries on all the examples we found in the algorithmic literature on that subject and on nearly all the examples we constructed.

We also experimented the heuristic on random maximal planar graphs (using the LEDA generator): As expected, most of such graphs with few vertices do have symmetries and there are very few as soon as the vertex number exceeds twenty. More precisely, the proportion of random maximal planar graphs that do have symmetries is roughly 95% (resp. 70%, 35%, 15%, 5%) when the number of vertices is 8 (resp. 10, 12, 14, 16).

The overall complexity of the heuristic is $\mathcal{O}(n^3)$, (i.e. the complexity of the computation of the eigenvalues). It certainly could be improved, but that was not our main goal. The different heuristics to identify the isomorphic subgraphs, compute the rotations etc are either linear or in $\mathcal{O}(n \log n)$. To give a rough idea of the time computation, on a Pentium running at 200MHz, finding a symmetry in a graph with 100 vertices takes 0.45 seconds and in a graph with 400 vertices 32 seconds. The code source of the C++ implementation is less than 1500 lines.

Acknowledgment

The author would like to express his gratitude to J.P. Benzécri who introduced him the Factorial Analysis techniques, to Jarik Nesětril and to P. Ossona de Mendez for his helpful comments and remarks.

References

- [1] J.P. Benzécri and et al, *L'analyse des données — Tome II: l'analyse des correspondances*, Dunod, Paris, 1973.
- [2] M. Capobianco, *Examples and counterexamples in graph theory*, Elsevier, North Holland, New-York, 1978.
- [3] C.N.R.S. (ed.), *L'analyse factorielle et ses applications*, Paris, 1956.
- [4] D.G. Corneil and C.C. Gotlieb, *An efficient algorithm for graph isomorphism*, Journal of the ACM **17** (1970), no. 1, 51–64.
- [5] H. de Fraysseix and P. Ossona de Mendez, *Discovering automorphisms using abstract distances*, Tech. report, KAM-DIMATIA Series, 1999, (in preparation).
- [6] H. de Fraysseix and Kuntz P., *Pagination of large scale networks*, Algorithms review **2** (1992), no. 3, 105–112.
- [7] I.S. Filotti and J.N. Mayer, *A polynomial-time algorithm for determining the isomorphism of graphs of fixed genus (working paper)*, Conference Proceedings of the Twelfth Annual ACM Symposium on Theory of Computing (Los Angeles, California), 1980, pp. 236–243.

- [8] S.L. Hakimi and S.S. Yau, *Distance matrix of a graph and its realizability*, Quarterly of Applied Mathematics **22** (1964), 305–317.
- [9] S-H. Hong, P. Eades, and S-H. Lee, *Finding planar geometric automorphisms in planar graphs*, ISAAC 98 (ed. Chwa and Ibarra), Springer Lecture Notes in Computer Science 1533 (1998), 277–286.
- [10] S-H. Hong, P. Eades, and S-H. Lee, *Drawing series parallel digraphs symmetrically* (to appear).
- [11] J.E. Hopcroft and J.K. Wong, *Linear time algorithm for isomorphism of planar graphs (preliminary report)*, Conference Record of Sixth Annual ACM Symposium on Theory of Computing (Seattle, Washington), 1974, pp. 172–184.
- [12] J.B. Kruskal and J.B. Seery, *Designing network diagrams*, Proceedings of the First General Conference on Social Graphics, 1978, pp. 22–50.
- [13] P. Kuntz, *Représentation euclidienne d'un graphe en vue de sa segmentation*, Ph.D. thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris, 1992.
- [14] G.S. Lueker and S.B. Kellogg, *A linear time algorithm for deciding interval graph isomorphism*, Journal of the ACM **26** (1979), no. 2, 183–195.
- [15] J. Manning, *Geometric symmetry in graphs*, Ph.D. thesis, Purdue University, New York, 1990.
- [16] J. Manning and M.J. Atallah, *Fast detection and display of symmetry in trees*, Congressus Numerantium **64** (1988), 159–169.
- [17] ———, *Fast detection and display of symmetry in outerplanar graphs*, Discrete Applied Mathematics **39** (1992), 13–35.
- [18] G.L. Miller, *Isomorphism testing for graphs of bounded genus*, Conference Proceedings of the Twelfth Annual ACM Symposium on Theory of Computing (Los Angeles, California), 1980, pp. 225–235.
- [19] ———, *Isomorphism of graphs which are pairwise k -separable*, Information and Control **56** (1983), no. 1/2, 21–33.
- [20] ———, *Isomorphism of k -contractible graphs. a generalization of bounded valence and bounded genus*, Information and Control **56** (1983), no. 1/2, 1–20.