────────────── **CORRESPONDENCE** ──────────────

Comment on "Isomorphism, Automorphism Partitioning, and Canonical Labeling Can Be Solved in Polynomial-Time for Molecular Graphs"

Bo Tao Fan, Annick Panaye, and Jean-Pierre Doucet*

Institut de Topologie et de Dynamique des Systèmes, CNRS UPRES-A 7086, Université Paris 7,
1, rue Guy de la Brosse, 75005 Paris, France

In a recent paper dealing with the search for equivalent sites,[1] Faulon reported a counter example graph (Figure 1) to the algorithm we proposed for the detection of the constitutionally equivalent sites.[2] Faulon concluded that our algorithm will fail for such a graph according to the statement "two equivalent atoms must have identical subspanning trees (MOST and MINST)". In fact our program gives correct and complete solution for this graph provided that the MOST and MINST technique is applied with a simple condition: for two comparative sites, their corresponding layers must have the same number of nonrepetitive nodes. If this condition cannot be satisfied, it is not necessary to perform next step.

In fact, the definition of MINST allows for the repetition of the same node number in the same layer (but the nodes which have been used cannot appear in next layer). This property is indispensable when we compare two sites, one example being shown in Figure 2. The nodes 2 and 4 of the example graph are constitutionally equivalent. If the repetitive use of a same node number in a same layer is prohibited, we have different ways to extract the MINSTs, one possibility being shown in (a), which leads us to conclude that the two nodes 2 and 4 are not equivalent. By contrast, in the case where the repetition of the same node is tolerated, we have only one possibility to extract the MINST (see Figure 2b).

This tolerance for extracting the MINST can bring about another problem: the possibility of extending the ring size from n to $n + 1$. In Figure 3, we can remark that the MINST for a four-membered ring has the same graph as that for a five-membered ring. To avoid this possible erroneous comparison, we must carry out the described test before the layer-by-layer comparison of two MINSTs. This test is illustrated in Figure 4 using Faulon's graph.

In this figure, for Faulon's graph, the vectors of numbers of nonrepetitive nodes are (3,5,5,2) for MINST(1), and (3,6,5,1) for MINST(8), respectively. The nonequivalence is immediately detected without comparing for each layer the corresponding nodes. In fact, this preliminary test excludes many nonequivalent sites before the comparison. By consequence, it could increase the speed of the process for certain structures. But for other graphs, the process could be more time-consuming because of the introduction of this new test. A comparative study by computing the examples used in our previous paper[2] shows that the average speed is
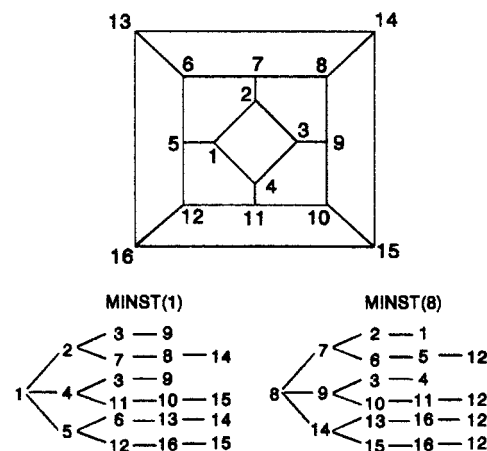

Figure 1. Faulon's graph. In this graph, sites 1 and 8 are not equivalent, but they possess identical minimum not overlapping spanning trees (MINST).
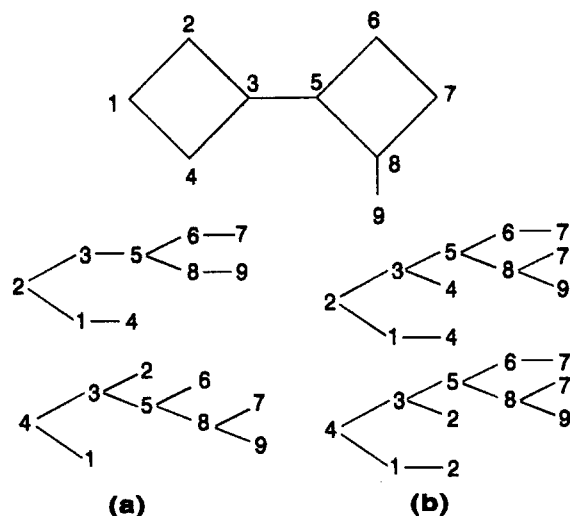

Figure 2. Indispensability of repetitive use of a node in the same layer of MINST: (a) without permission of repetition, two equivalent sites could be missed and (b) with permission of repetition, two equivalent sites are determined.

practically not changed. We thank Dr. Faulon for his example which shows again the importance of this test, specially in graphs containing rings the size of which only differs by one bond and where all vertices have the same degree.
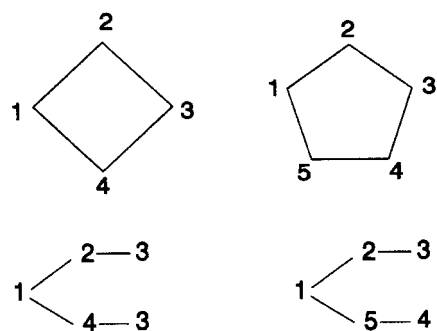
CORRESPONDENCE

J. Chem. Inf. Comput. Sci., Vol. 39, No. 3, 1999 **631**

Figure 3. Problem of extending the ring size caused by repetitive use of a node in the same layer of MINST.

As pointed out by Faulon, our algorithm can be very efficient, because it is polynomial. In fact, in the worst case, the estimation of operation complexity can be made by taking into account different steps. The first step concerns the categorization using identifier values with maximum $N/2$ iterations (N means the total number of nodes). It results in maximum operation numbers $N^2/2$, in addition to the assignment of initial identifier values (N operations). This is followed by a comparison of the identifier values comprising $N(N-1)/2$ operations. The next step is the detection of possible identical MINST which comprises of two substeps: extraction of subspanning trees from connection table and comparison of the subspanning trees. In the worst case where all identifier values are different, the computational complexities for two substeps are both $O(N^2)$ [3]. The last step,
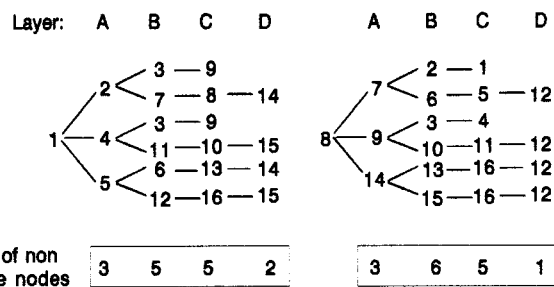

Figure 4. Preliminary test on the number of nonrepetitive sites for two corresponding sites.

refining the partitioning, is the repetition of the comparison step of subspanning tree, its complexity is also $O(N^2)$. Finally, the computational complexity for our algorithm can be estimated as $O(N^2)$.

REFERENCES AND NOTES

(1) Faulon, J. L. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 432−444.

(2) Fan, B. T.; Barbu, A.; Panaye, A.; Doucet, J. P. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 654−659.

(3) Gaudel, M. C.; Soria, M.; Froidevaux, C. Types de données et algorithmes − Recherche, tri, algorithmes sur les graphes, INRIA ed.; 1992; Vol. II, France.