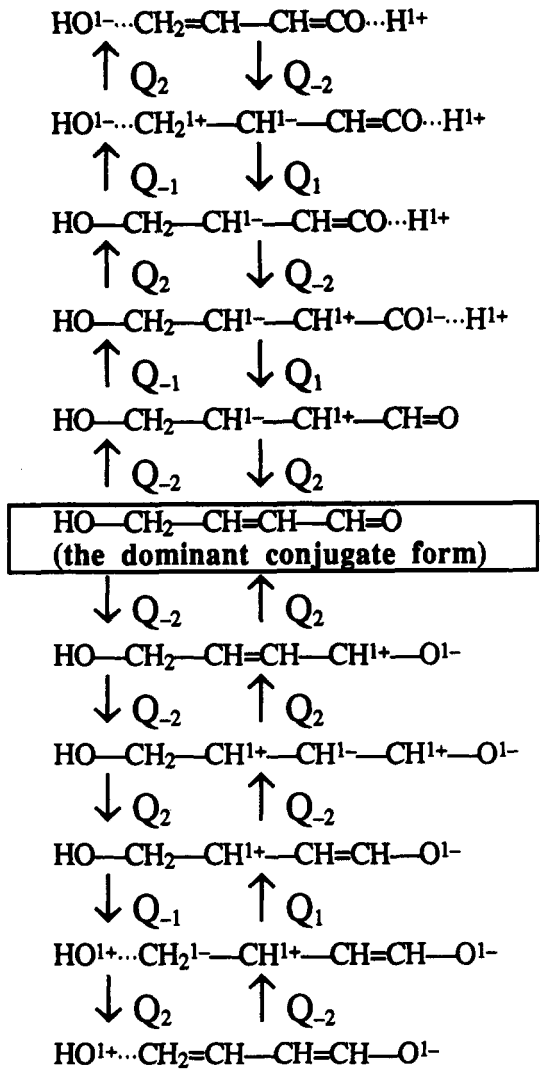

Figure 1. Formation of conjugates through bond-forming operators, Q_i, and bond-dissociating operators, Q_{-i}.

neering Research Center of the University of Maryland; by the Systems Research Center funded by the National Science Foundation (NSFD CDR 8803012); and by a Research Initiation Award from the National Science Foundation (CTS 9010549).

Registry No. Pentane, 109-66-0; 2-methylbutane, 78-78-4; 2,2-dimethylpropane, 463-82-1; 2-methylheptane, 592-27-8; 3-methylheptane, 589-81-1; 4-methylheptane, 589-53-7; 1-pentene, 109-67-1; *cis*-2-pentene, 627-20-3; *trans*-2-pentene, 646-04-8; 2-methyl-1-butene, 563-46-2; 2-methyl-2-butene, 513-35-9; ethane, 74-84-0; propane, 74-98-6; butane, 106-97-8; isobutane, 75-28-5; hexane, 110-54-3; 2-methylpentane, 107-83-5; 3-methylpentane, 96-14-0; 2,2-dimethylbutane, 75-83-2; 2,3-dimethylbutane, 79-29-8; heptane, 142-82-5; 2-methylhexane, 591-76-4; 3-methylhexane, 589-34-4; 2,2-dimethylpentane, 590-35-2; 2,3-dimethylpentane, 565-59-3; 2,4-dimethylpentane, 108-08-7; 3,3-dimethylpentane, 562-49-2; 3-ethylpentane, 617-78-7; 2,2,3-trimethylbutane, 464-06-2; 2,2-dimethylhexane, 590-73-8; 2,3-dimethylhexane, 584-94-1; 2,4-dimethylhexane, 589-43-5; 2,5-dimethylhexane, 592-13-2; 3,3-dimethylhexane, 563-16-6; 3,4-dimethylhexane, 583-48-2; 3-ethylhexane, 619-99-8; 2,2,3-trimethylpentane, 564-02-3; 2,2,4-trimethylpentane, 540-84-1; 2,3,3-trimethylpentane, 560-21-4; 2,3,4-trimethylpentane, 565-75-3; 2-methyl-3-ethylpentane, 609-26-7; 3-methyl-3-ethylpentane, 1067-08-9; nonane, 111-84-2; 2-methyloctane, 3221-61-2; 2,2-dimethylheptane, 1071-26-7; 2,2,3-trimethylhexane, 16747-25-4; 2,2,4-trimethylhexane, 16747-26-5; 2,2,5-trimethylhexane, 3522-94-9; 3,3-diethylpentane, 1067-20-5; 2,2,3,3-tetramethylpentane, 7154-79-2; 2,2,3,4-tetramethylpentane, 1186-53-4; 2,2,4,4-tetramethylpentane, 1070-87-7; 2,3,3,4-tetramethylpentane, 16747-38-9; decane, 124-18-5; 3,3,5-trimethylheptane, 7154-80-5; octane, 111-65-9.

REFERENCES AND NOTES

(1) Reid, R. C.; Prausnitz, J. M.; Poling, B. E. *The Properties of Gases and Liquids*, 4th ed.; McGraw-Hill: New York, 1987.
(2) Cox, J. D.; Pilcher, G. *Thermochemistry of Organic and Organometallic Compounds*; Academic Press: New York, 1970.
(3) Joback, K. G.; Reid, R. C. Estimation of Pure-Component Properties from Group Contributions. *Chem. Eng. Commun.* **1987**, *57*, 233-243.
(4) Benson, S. W. *Thermochemical Kinetics*; Wiley: New York, 1968.
(5) Domalski, E. S.; Hearing, E. D. Estimation of the Thermodynamic Properties of Hydrocarbons at 298.15 K. *J. Phys. Chem. Ref. Data* **1988**, *14*, 1637-1678.
(6) Mavrovouniotis, M. L.; Bayol, P.; Lam, T.-K. M.; Stephanopoulos, G.; Stephanopoulos, G. A Group-Containing Method for the Estimation of Equilibrium Constants of Biochemical Reactions. *Biotechnol. Tech.* **1988**, *2*, 23-28.
(7) Mavrovouniotis, M. L. Group Contributions to the Gibbs Energy of Formation of Biochemical Compounds in Aqueous Solution. *Biotechnol. Bioeng.* **1990**, *36*, 1070-1082.
(8) Morrison, R. T.; Boyd, R. N. *Organic Chemistry*, 3rd ed.; Allyn and Bacon: Boston, 1973.
(9) Fliszar, S. *Charge Distributions and Chemical Effects*; Springer-Verlag: New York, 1983.
(10) Mullay, J. A Simple Method for Calculating Atomic Charges in Molecules. *J. Am. Chem. Soc.* **1986**, *108*, 1770-1775.
(11) Mavrovouniotis, M. L. Estimation of Properties from Conjugate Forms of Molecular Structures: The ABC Approach. *Ind. Eng. Chem. Res.* **1990**, *29*, 1943-1953.

On Using the Adjacency Matrix Power Method for Perception of Symmetry and for Isomorphism Testing of Highly Intricate Graphs

GERTA RÜCKER and CHRISTOPH RÜCKER*
Institut für Organische Chemie und Biochemie, Universität Freiburg, Albertstrasse 21, D-7800 Freiburg, FRG

Received August 13, 1990

A modification of the adjacency matrix power method described recently for the perception of symmetry in graphs is introduced, which expands the limits of the method far beyond the realm of chemically interesting graphs. The procedure finds the automorphism partition even for intricate graphs without performing a tree search. The calculation effort increases with the problem size polynomially for all tested cases, including strongly regular graphs, two-level regular graphs, and graphs corresponding to balanced incomplete block designs (BIBD). An equally powerful computer program for testing isomorphism of graphs based on the adjacency matrix power method is introduced.

Recently we described a computer program (TOPSYM) for the perception of constitutional (topological) symmetry in molecular graphs which is based on a very elementary principle, raising the adjacency matrix to its higher powers,

0095-2338/91/1631-0123$02.50/0

a)

1	11	12	13	14	15	16
2	11	12	17	18	19	20
3	11	13	17	21	22	23
4	11	14	18	21	24	25
5	12	13	19	22	24	25
6	12	15	20	21	23	24
7	13	16	18	20	23	25
8	14	15	17	19	23	25
9	14	16	17	20	22	24
10	15	16	18	19	21	22

b)

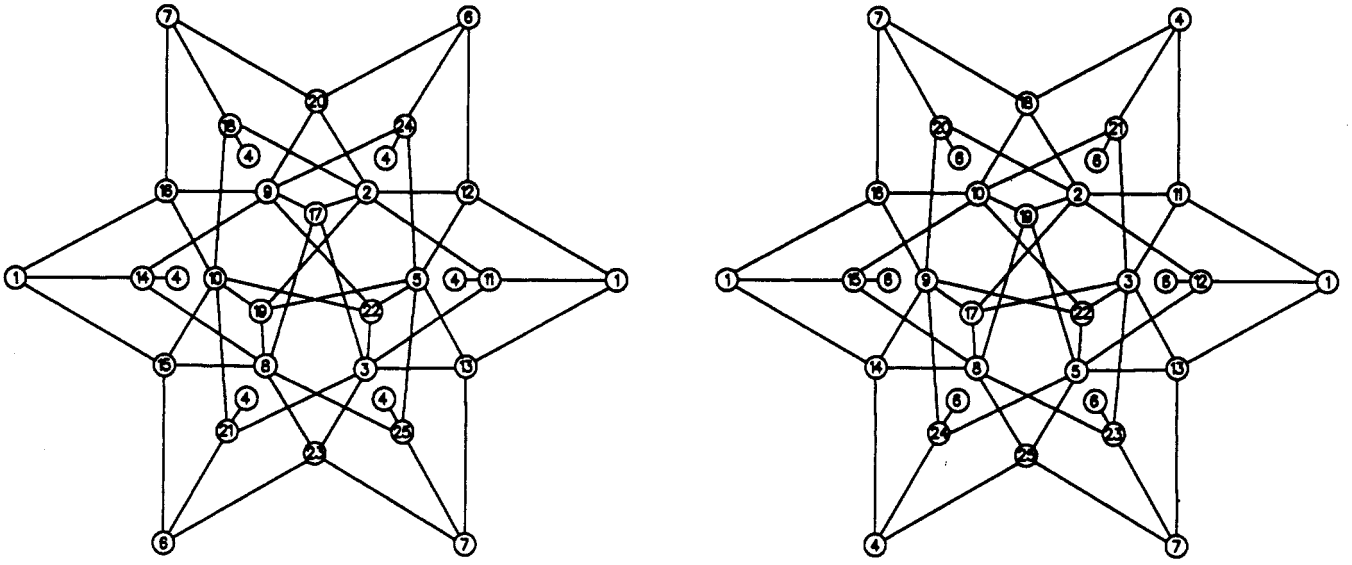

Figure 1. Mathon's graph A_{25}. (a) Connection table. (b) Two out of several equivalent graphical representations. In each such drawing three of the four equivalent vertices 1,4,6,7 appear twice and one appears six times in order to keep the figure as clear as possible.

evaluating the entries in these higher matrices, and thus partitioning the atoms and pairs of atoms into equivalence classes.[1a] Though the method works satisfactorily for all molecular graphs tested,[1b] we did not claim the method to be useful for *any* graph.

More than a decade ago, R. Mathon published a collection of intricate and highly artificial ("mathematical") graphs for the purpose of challenging algorithms for graph isomorphism testing.[2] Some of these graphs are strongly regular, some are two-level regular, others correspond to balanced incomplete block designs. They all are very unusual (at least in a chemist's view), and, as Davis and Ellzey[3] state "it is highly unlikely that one would encounter them or others like them in any practical application of group theoretical techniques". These authors,[3] surprisingly, seem to be the only ones to accept Mathon's challenge at least partially. The aim of the present study was to expand our simple method[1a] until it is able to treat the Mathon graphs correctly.

Figure 1 shows different graphical representations of the first graph in Mathon's collection, A_{25}. No fully satisfying representation of this graph seems to exist, and any correct representation is misleading in some respect. Note that these drawings were constructed using the published automorphism partition, they are therefore of no help in finding this partition. This example is given here to emphasize that perceiving the symmetry and testing the identity of graphs is not a trivial task.

A characteristic feature common to all the Mathon graphs is that in their adjacency matrices and in *all* the higher powers thereof both the diagonal elements and the patterns of entries in the off-diagonal elements are indistinguishable for some nonequivalent vertices (or, in many cases, even for *all* vertices, without any two vertices being equivalent). This means that the differences between nonequivalent vertices are so subtle that the simple TOPSYM cannot perceive them.

A straightforward modification of the program, however, enables it to treat the Mathon graphs (class I and II) correctly. The key to the successful analysis of these graphs is the observation that marking different pairs of vertices[4] (e.g., by attaching substituents or by introducing an additional edge) may result in derivative graphs which differ in an easily obtained graph invariant such as the number of classes of vertices. Such graphs are easily perceived as nonisomorphic by inspection or by TOPSYM. Two vertex pairs for which this happens are obviously nonequivalent, and one has now to conclude in a legitimate manner from the nonequivalence of pairs to the nonequivalence of single vertices.[5] The modified TOPSYM (MATSYM2) does exactly this: Given a graph of n vertices it accumulates information on all the graphs derived from the original one by introducing an additional edge connecting each particular pair of vertices i and j, one after the other. The program treats each such graph in the normal TOPSYM manner and writes a graph invariant (the number of classes of vertices[6]) into the appropriate element i,j of a new $n \times n$ matrix. After this matrix is completely filled, the program looks up patterns of entries in its columns or rows.[7] When the patterns are different for any two vertices i and j of the original graph, i and j are nonequivalent, they are now marked as such, and a final TOPSYM run[8] is done for the original graph using this nonequivalence information from the beginning. *The resulting vertex partition is an exact reproduction of the published automorphism partition for all the Mathon class I and II graphs as well as all the Weisfeiler*[9] *graphs.* The price for the increased discriminating power has to be paid, of course, in CPU time: Instead of one TOPSYM run about $n^2/2$ (simplified) TOPSYM runs are required.

For the Mathon class III graphs, which are two-level regular, the method is still not powerful enough. These are, however, successfully analyzed by a further logical extension

Table I. Results of MATSYM2 Treatment of Mathon Class I and II and Weisfeiler Graphs

structure (graph)	n	classes of vertices	classes of pairs	CPU time[a] MATSYM2
	(a) Mathon Class I Graphs			
A_{25}	25	4	19	11.87
B_{25}	25	5	24	11.84
A_{50}	50	50	1225	604.96
B_{50}	50	50	1225	699.86
C_{50}	50	50	1225	699.23
D_{50}	50	50	1225	697.51
	(b) Mathon Class II Graphs			
A^1_{25}	25	8	62	11.61
B^1_{25}	25	25	300	11.62
A^1_{29}	29	29	406	25.94
A^3_{29}	29	29	406	25.96
B^1_{29}	29	29	406	25.99
B^3_{29}	29	29	406	25.93
B^5_{29}	29	29	406	25.89
B^7_{29}	29	29	406	25.93
B^9_{29}	29	29	406	25.89
B^{11}_{29}	29	29	406	25.93
A_{35}	35	35	595	69.03
B_{35}	35	35	595	69.07
C_{35}	35	35	595	69.02
D_{35}	35	35	595	69.13
	(c) Weisfeiler Graphs			
W1031	10	1	2	0.13
W1361	13	1	2	0.41
W1561	15	1	2	0.80
W1651	16	1	2	1.14
W1661	16	1	2	1.17
W1662	16	1	3	1.16
W1781	17	1	2	1.57
W21101	21	1	2	4.88
W2581	25	1	2	12.09
W25121	25	15	160	11.67
W25122	25	15	160	11.77
W25123	25	25	300	11.58
W25124	25	25	300	11.57
W25125	25	9	100	11.63
W25126	25	7	60	11.65
W25127	25	15	160	11.74
W25128	25	9	100	11.62
W25129	25	7	60	11.61
W251210	25	8	62	11.62
W251211	25	3	11	11.60
W251212	25	8	62	11.61
W251213	25	15	160	11.63
W251214	25	3	11	11.64
W251215	25	1	2	11.64
W26101	26	10	97	14.87
W26102	26	9	70	14.94
W26103	26	8	67	14.86
W26104	26	8	67	14.86
W26105	26	16	175	14.89
W26106	26	10	109	14.87
W26107	26	2	9	14.85
W26108	26	2	8	14.91
W26109	26	26	325	14.82
W261010	26	16	175	14.97
W27101	27	1	2	17.43
W28121	28	2	10	21.89
W28122	28	2	7	21.85
W28123	28	2	8	21.86
W28124	28	1	2	22.09

[a] In seconds, IBM 3090.

of the procedure, MATSYM3: All derivative graphs are constructed in which triples of vertices i,j,k are marked by introducing three additional edges i–j, j–k, and i–k. For every such graph the number of classes of vertices found in a TOPSYM run is written into a new $n \times n \times n$ matrix, and the patterns of entries in this matrix are evaluated for discrimination between vertices. Another TOPSYM run[8] again finds the final vertex partition and vertex pair partition for the original graph. This procedure, of course, is even more time consuming, requiring about $n^3/6$ (simplified) TOPSYM runs.

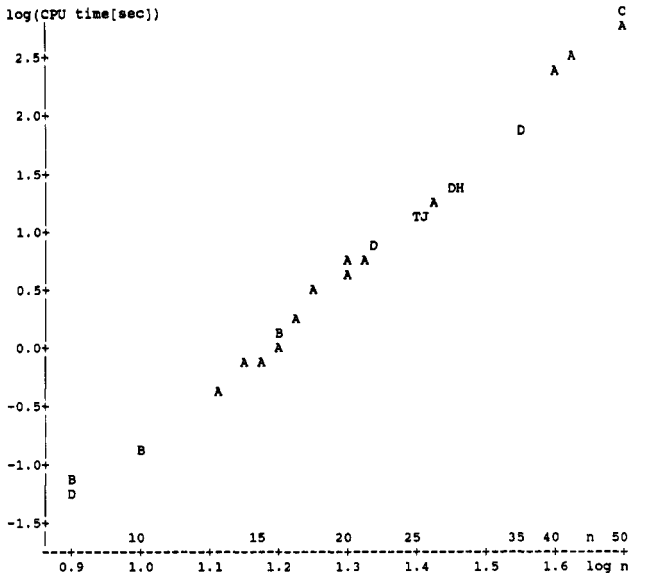

Figure 2. Plot of log [CPU time (s)] vs log (n) for MATSYM2 treatment of Weisfeiler, Mathon class I and II, and some other graphs. A = 1 observation, B = 2 observations, etc.

Table II. Results of MATSYM3 Treatment of Mathon Class III Graphs

structure (graph)	n	classes of vertices	classes of pairs	CPU time[a] MATSYM3
A_{52}	52	4	32	10691
B_{52}	52	8	142	10653
A_{60}	60	8	170	27037
B_{60}	60	18	480	27036

Vertex Partitions

A_{52} (1,4,5,12,13,14,21,22,23,24,25,26,27,30,31,38,39,40,47,48,49, 50,51,52) (2,28) (3,29) (6,7,8,9,10,11,15,16,17,18,19,20,32, 33,34,35,36,37,41,42,43,44,45,46)

B_{52} (1,5,6,14,19,23,27,31,32,40,45,49) (2,9,26,28,35,52) (3,7,15,20,22,25,29,33,41,46,48,51) (4,30) (8,13,18,34,39,44) (10,16,17,36,42,43) (11,12,21,37,38,47) (24,50)

A_{60} (1,2,3,15,22,25,31,32,33,45,52,55) (4,34) (5,7,35,37) (6,8,36,38) (9,39) (10,14,17,21,26,30,40,44,47,51,56,60) (11,12,18,19,28, 29,41,42,48,49,58,59) (13,16,20,23,24,27,43,46,50,53,54,57)

B_{60} (1,4,31,34) (2,5,32,35) (3,6,33,36) (7,24,37,54) (8,23,38,53) (9,16,39,46) (10,25,40,55) (11,26,41,56) (12,22,42,52) (13,21,43,51) (14,44) (15,45) (17,27,47,57) (18,28,48,58) (19,49) (20,50) (29,59) (30,60)

[a] In seconds, IBM 3090.

Though we were able to analyze all the graphs tried by using one of our procedures, again we do not claim that *any* graph can successfully be treated by our programs. On the contrary, it is to be expected that even more demanding graphs can be constructed by group theoretical or combinatorial methods, and the nontransitive three-level regular graph of 139 300 vertices mentioned by Mathon is an obvious candidate.

Since the program does not perform a tree search with its notoriously exponential dependence on the problem size, it was expected to exhibit polynomial complexity. This is confirmed by the experimental results. Table I gives the numbers of vertex classes, of vertex pair classes, and the CPU times (in seconds) (on a IBM 3090 computer) for MATSYM2 treatment of the Mathon class I and II and the Weisfeiler graphs. Figure 2 is a log–log plot of the CPU times vs n (another few graphs are included). From the slope of the least-squares straight line ($r^2 = 0.9908$) obtained in this plot, a value of 5 for the exponent in the polynomial dependence is calculated. Table II gives the corresponding data for MATSYM3 treatment of the Mathon class III graphs. Since the automorphism partitions of these graphs are not published, the partitions found by

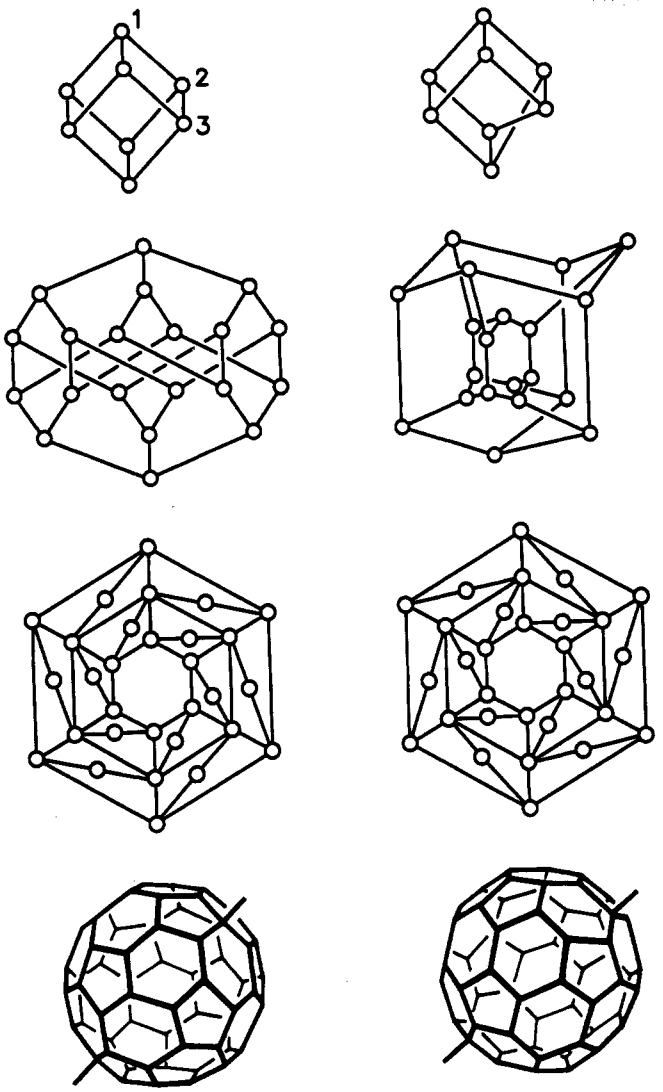

Figure 3. Pairs of graphs which were tested for isomorphism by ISOTOP. The CPU times elapsed in the run are given. From top to bottom: Cubane and Möbius cubane, not isomorphic, 0.020 s; two representations of the Desargues–Levi graph, isomorphic, 0.276 s; the two Fisher graphs, not isomorphic, 0.378 s; two disubstituted soccerane derivatives, not isomorphic, 7.981 s.

MATSYM3 are given here. Again a log–log plot of the CPU times vs n (including another few graphs) yields a straight line, the slope of which gives a value of 6 for the exponent.[10]

Having in our hands a powerful program for detecting the symmetry of graphs, i.e., finding the automorphism partition, we could not resist the temptation to use this tool for graph isomorphism testing. Testing the isomorphism of two graphs G and G' of n vertices each is equivalent to finding the automorphism partition of their disjoint union, that is of a graph of $2n$ vertices.[11,12] If in every vertex class of the union graph there is for each vertex in G a corresponding vertex in G', then the graphs are regarded as isomorphic.

In fact, for chemical graphs a modification of TOPSYM proved sufficient (ISOTOP), correctly finding, e.g., the nonidentity of cubane and Möbius cubane[13] (see Figure 3), the identity of different representations of the Desargues–Levi graph,[14] and the nonidentity of the two Fisher graphs[14] or of the two disubstituted derivatives of "soccerane" shown.

Correspondingly, testing the isomorphism of the Mathon graphs requires a modification of MATSYM2 (ISOMAT2). This program, for example, correctly perceived the identity of two of the Mathon graphs with two of the Weisfeiler graphs[15] (A^1_{25} = W251210, 99.22 s; B^1_{25} = W25123, 98.94 s) and the pairwise nonidentity of all the eight strongly regular Mathon graphs which consist of 29 nonequivalent vertices of degree 14 each (A^1_{29} – B^{11}_{29}, 227 s for each comparison). Likewise, the four Mathon graphs of 35 vertices and the four Mathon graphs of 50 vertices were shown to be pairwise nonisomorphic.

The dependence of the CPU time on the problem size for the isomorphism programs is again polynomial, with an exponent between 3 and 4 for ISOTOP and 5 for ISOMAT2.

Since our symmetry perception programs provide a partition of the pairs of vertices as well as of the vertices themselves into equivalence classes, otherwise similar graphs are often easily seen to be nonidentical (without the need to employ an isomorphism testing program) by differences in the number or sizes of these classes. Thus, out of the four strongly regular Weisfeiler graphs of 28 vertices of degree 12 (see Table I), three do not differ in the number of vertex classes but do differ in the number of vertex pair classes. Perceiving these four graphs as all nonidentical therefore does not require six time-consuming ISOMAT2 runs (186 s of CPU time each) but requires just the four MATSYM2 runs, 88 s altogether.

Thus, our conceptually simple and, as far as we know, new approach to the graph automorphism and isomorphism problems compares favorably with existing ones.[3,11,12,16,17] After having successfully tested our algorithms with the demanding Mathon graphs, we feel rather confident that they will be useful and reliable in the field of chemical graphs. A copy of the programs is available upon request from the authors.

REFERENCES AND NOTES

(1) (a) Rücker, G.; Rücker, C. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 187. (b) E.g., for a four-rung Möbius latter molecule (homeomorphic to Möbius cubane, see 2nd graph in Figure 3) the number of vertex classes found by TOPSYM is identical with one obtained by a thorough mathematical analysis and published recently, see: Walba, D. M.; Simon, J.; Harary, F. *Tetrahedron Lett.* **1988**, *29*, 731.
(2) Mathon, R. *Proc. 9th S.-E. Conf. Combinatorics, Graph Theory and Computing* **1978**, 499.
(3) Davis, M. I.; Ellzey, M. L., Jr. *J. Comput. Chem.* **1983**, *4*, 267.
(4) Marking single vertices does not suffice.
(5) The word "legitimate" means the following. Consider, for example, the graph corresponding to the $(CH)_8$ hydrocarbon cubane (Figure 3). Vertex pairs 1,2 and 1,3 are of course different. Nevertheless, it would be illegitimate to conclude that vertices 2 and 3 are constitutionally different. The legitimate procedure described below avoids this pitfall.
(6) In principle *any* graph invariant can be used for this purpose. The number of vertex classes is chosen since it is both readily obtainable (more readily than, e.g., the number of classes of vertex pairs) and of sufficiently high discriminating power (in contrast to, e.g., the Wiener number).
(7) Since the simple TOPSYM fails to find the correct number of vertex classes for the original Mathon graph, we have to realize the possibility that the same will happen for at least some of the derivative graphs. Therefore some elements of this new matrix will contain "wrong" information, as judged by the correct number of vertex classes for the respective derivative graph. The point is, however, that the physical meaning of these matrix elements is unimportant. The only conditions to be met by these matrix elements are as follows: (a) The matrix elements have to be the same for every two pairs i,j and i',j' which are in fact equivalent. (b) The matrix elements have to be different for a sufficiently high number of nonequivalent pairs i,j and i',j'. Condition a is met since TOPSYM treats like things like. Condition b is met at least in all Mathon class I and II graphs and in all the Weisfeiler[9] graphs.
(8) The TOPSYM procedure has been further improved, to be described elsewhere.
(9) Weisfeiler, B. *On Construction and Identification of Graphs*; Lecture Notes in Mathematics No. 558; Springer: Berlin, 1976.
(10) It has now become clear that the corresponding exponent for the simple TOPSYM[1a] is between 3 and 4.
(11) Read, R. C.; Corneil, D. G. *J. Graph Theory* **1977**, *1*, 339. Gati, G. *J. Graph Theory* **1979**, *3*, 95.
(12) Fowler, G.; Haralick, R.; Gray, F. G.; Feustel, C.; Grinstead, C. *Artif. Intell.* **1983**, *21*, 245.
(13) Randić, M. *J. Comput. Chem.* **1980**, *1*, 386.
(14) Randić, M. *SIAM J. Alg. Disc. Methods* **1985**, *6*, 145.
(15) Walter, C. D. *SIAM J. Alg. Disc. Methods* **1986**, *7*, 18.
(16) (a) Balaban, A. T.; Mekenyan, O.; Bonchev, D. *J. Comput. Chem.* **1985**, *6*, 538; (b) Bonchev, D.; Mekenyan, O.; Balaban, A. T. *MATCH* **1985**, *18*, 83.
(17) Stinson, D. R. *Ars Combinatoria* **1985**, *19*, 213.