Acta Cryst. (1978). A**34**, 275–282

A Graph Theoretical Basis for Structural Chemistry. I. Structures Based on Trivalent Graphs with 10 Vertices

BY MILAN RANDIĆ

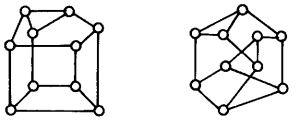
The Ames Laboratory–USDOE, Iowa State University, Ames, Iowa 50011, *USA*

(*Received* 3 *November* 1976; *accepted* 12 *October* 1977)

This paper is the first in a series concerned with a particular graph theoretical scheme for enumeration and derivation of structures of prescribed form. The problem spoken to by this work is that of finding a unique formal procedure for generating all distinct (*i.e.* non-isomorphic) graphs of a given number of vertices and prescribed valency. Here the scheme is outlined as applied to a search for regular trivalent graphs, some of which correspond to trivalent polyhedra. A systematic procedure for obtaining the number of equivalence classes of the adjacency matrices associated with trivalent graphs of n points is described. The procedure is of general applicability, though no proof of its correctness is offered. Instead, a number of examples are discussed, and its application is illustrated. The scheme is based on consideration of unique matrices associated with graphs which in turn are determined so that the corresponding binary code obtained by reading the rows of the matrix from left to right and from top to bottom represents the smallest possible binary code. Part of the scheme consists in finding all acceptable matrices, testing them for isomorphism, and selecting those which satisfy additional restrictions and ensure that the derived graph represents a structure of interest.

Introduction

Graph theory is concerned, among other problems associated with connectivity, with enumeration and construction of structures of prescribed form. As the number of atoms in a structure increases, omissions and duplications are more likely; great care must therefore be exercised to avoid these errors. The pictorial representations of structures may appear distinctive when in fact the structures are not. In order to eliminate any possible duplicates, one needs a reliable test for graph isomorphism. Lack of a suitable procedure has plagued many previous applications, leaving some uncertainty about the reported data. For example, the number of trivalent polyhedra with $n = 18$ vertices was believed to be 1250 and has been revised to 1249, but even the revised figure, derived by use of a computer, is uncertain.* The problem of recognizing identical graphs is not simple even for relatively small graphs. For example the two trivalent graphs with $n = 10$ vertices



initially appeared as distinctive, but were soon recognized as identical (Balaban, 1966, 1967). One can therefore anticipate difficulties in recognizing isomorphs when the number of vertices in a structure increases. These difficulties may account for the fact that no list of trivalent graphs with $n = 12$ vertices has been reported, despite the fact that their number (85) is by no means excessive.* Errors due to omissions are even more troublesome to correct than errors due to duplication because they are the result of faults of the algorithm adopted in the search. Hence, one needs some strong insurance that a method is free of such limitations, and, in addition, one needs a simple test of isomorphism to verify the results. Below we outline a general procedure for enumeration and construction of structures which satisfies the above requirements.

Outline of the method

The basis of our procedure is unique adjacency matrices which can each be assigned to specific graphs; these matrices follow from a canonical labeling of vertices. The form of the adjacency matrix of a graph depends on the assumed labels of the vertices. The adjacency matrix is an $n \times n$ matrix whose entries are either ones or zeros, depending on whether the pair of vertices is connected (1) or not (0). There are $n!$ different labelings; the number of distinctive matrices, however, will depend on the presence of equivalent vertices. Among different matrices one can single out

* The value 1250 was reported by Brückner (1900); the revised figure of 1249 is due to Grace (1965). For additional comments see Lederberg (1966).

* Balaban (1972) gives a list of 32 *planar* trivalent graphs only. A partial list of trivalent graphs (including multigraphs but excluding 1-connected graphs) has been also given by Lederberg (1965).

the matrix which corresponds to the smallest binary code when the entries are read from left to right and from top to bottom and use it as a unique descriptor (Randić, 1974). For the procedure to be practical, one must show that the unique matrix can be found without screening $n!$ possibilities. An algorithm that can accomplish the task has been suggested. The algorithm is essentially based on a search for labels in an *unlabeled* graph that ensures for each row of the matrix in succession that it correspond to the smallest binary code (Randić, 1977).* We will illustrate the derivation of labeling which produces a unique matrix on a convex polyhedron with eight vertices (Fig. 1). In order to reduce labor, we will use easily recognizable symmetry properties of the graph. An inspection of the pictorial diagram reveals only three non-equivalent vertices (shown as *a*, *b*, and *c* in Fig. 1). The label 1 can be assigned to any of the three non-equivalent sites, as shown at the top of Fig. 1. Only one of the three possibilities will later qualify for the smallest binary label, but at this initial stage we must consider all conceivable alternatives. Each of the initial cases will in a general situation lead to 3! possibilities for assignments of the largest labels 6, 7, and 8. For the first alternative we show three different assignments of the largest labels; the remaining three, because of the lateral symmetry of the figure, can be simply derived by exchanging left with right. Notice that in each case we generated two vertices already having two assigned labels. Hence the next smallest label (2) requires only a single new large label (5), but in order to have the second row of the adjacency matrix represent the smallest binary code from the available sites, we must select one in which the already assigned labels belong to the largest possible alternative – since this ensures the smallest number as a code. It follows therefore that label 2 should have as its neighbors 5,7,8 rather than 5,6,8 or 5,6,7. Only two of the three considered cases have the desired environment and require further pursuit. Of the two alternatives so derived we see immediately that only the first possibility will lead to the smaller code for the third row, giving for neighbors of label 3 the numbers 4,6,8; the second possibility will result in neighbors 4,6,7 and thus produce a larger binary code. In this way we have arrived in a few steps at the best possible labeling that follows the particular assignment of label 1. To complete the search, all other alternatives that can arise by placing label 1 in other non-equivalent sites must be fully examined. However, we immediately see that the alternative sites cannot produce a smaller code. For the second alternative case, we would have for label 1 a site with adjacent labels 7 and 8 which would make the second row as small as in the first case, but would make the third row correspond to a larger binary code, since label 3 cannot be adjacent to both 8 and 6, as was possible in the first case. The last alternative for label 1 makes the second row larger, as the label 2 must have two newly assigned labels. In order to appreciate the efficiency of the algorithm one should observe how many permutational possibilities have been discarded at each successive step: we have examined here fewer than a dozen possibilities from over 40 000, and even those under consideration when found unproductive have been abrogated at the earliest stage possible.

For our use here it is important that the procedure uniquely assign a matrix to a graph so that any two graphs can be tested for isomorphism in a relatively straightforward way. An alternative viewing of the isomorphism problem is that of *ordering* given graphs in a sequence. Obviously if two graphs fall in the same place in the sequence they are isomorphic. The ordering which we consider is of a lexicographic nature: each unique matrix represents a binary number, if interpreted by reading its rows from left to right and from top to bottom. Such numbers can be sequenced in an increasing order. Since any $n \times n$ matrix has a place somewhere in the sequence, if we can generate *all* such

* A brief account of the algorithm is to be found in Randić (1974). It should not be confused with an alternative algorithm discussed there at some length, which is based on permutation of rows and columns of an already *labeled* graph. It has been found since that such an approach does not always converge to an absolute minimum (Mackay, 1975).

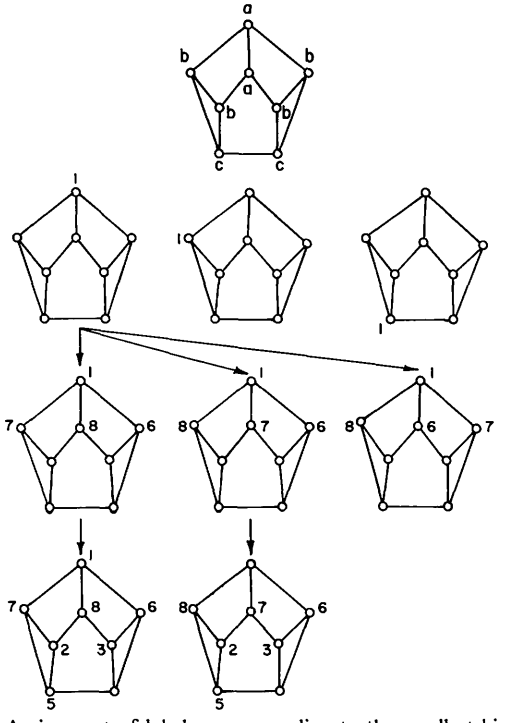


Fig. 1. Assignment of labels corresponding to the smallest binary code for an eight-vertices trivalent graph. The possibilities arising from permutations of labels among equivalent vertices are suppressed.

matrices we ensure that no omission will occur. By restricting our selection to those matrices that correspond to the smallest binary code we only eliminate duplicates.

In our constructions the only criterion used is that which requires the labels to give the smallest possible binary code: a structure is rejected if another alternative in the process of construction already defines the same connectivity. This guarantees that no omissions will result, and duplicates can be recognized by testing whether a derived labeling complies with the request for a minimal binary label. The above is, then, the essence of the procedure which we will apply to construction of all trivalent graphs with 10 and fewer vertices in order to illustrate the practicality of the approach.

Construction of graphs of particular form is of interest in various problems of chemistry and physics, and numerous specific applications have been considered. In connection with studies of valence isomerism of conjugated hydrocarbons Balaban (1970, 1972, 1973, 1974) has worked out many examples. In his approach Balaban used pictorial and geometrical arguments. In contrast, our scheme is of an analytical nature, based on manipulation of matrices, and hence is suitable to computer processing – an important advantage.

Trivalent graphs with six vertices

Before discussing the question of trivalent graphs with six vertices, let us point out that the smallest trivalent graph has four vertices; there is only one such graph, and only one matrix that satisfies the requirement of having in each row and column three non-zero entries:

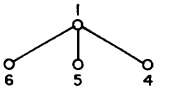
$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

The graph corresponds to a tetrahedron.

The search for trivalent graphs with six vertices we now view as a search for all six by six matrices which are symmetric, have zero on the diagonal, and have three non-zero entries in each row and column. Furthermore, each matrix must correspond to the smallest possible binary code for the structure, which of course eliminates an enormous number of undesired possibilities from consideration. Obviously the first row in all acceptable matrices has to be

0 0 0 1 1 1

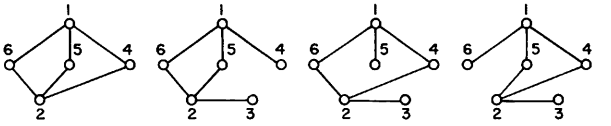
as this is the only smallest number consistent with the valency requirement. The above first row allows us to construct a part of the graphs sought; it shows that vertex (1) has three neighbors, (6), (5), and (4):



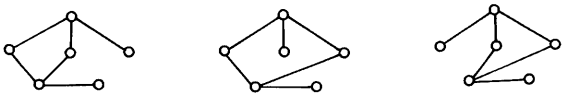
Now we proceed to consider acceptable forms for the second row of adjacency matrices. Since the matrix must be symmetric and have zero at the diagonal, we consider only four-digit binary numbers. This results in four distinctive possibilities, which, combined with the already established form for the first row, are shown below:

0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1
0 0 0 1 1 1 0 0 1 0 1 1 0 0 1 1 0 1 0 0 1 1 1 0.

Each of the partially derived adjacency matrix defines a fragment of a graph:



We see immediately that only two non-isomorphic fragments are produced and that the last two cases represent a duplication. Hence, in the continuation of the search they are excluded from consideration. Here, we eliminated the duplications by inspecting the pictures, which is not an analytical procedure; let us also explain how this isomorphism can be recognized from the partial matrices. In the following we pretend that we are not aware that the three last graphs are isomorphic. From the available partial matrices we can construct partial graphs, which we will, however, leave *unlabeled*:



We will consider each of the three graphs separately and will treat them as completed. The test for isomorphism consists in a search for the unique labels. In Fig. 2 the search is outlined for the first of the three graphs. Obviously label 1 belongs to vertices of the lowest valency, which in this case are terminal vertices. Their adjacent vertex must then have label 6, which gives for the first row of the matrices sought:

0 0 0 0 0 1.

There are two such terminal vertices; hence, two assignments need to be further considered. Each of the two possibilities predetermines the site for label 2 for the remaining terminal vertex, which has as its neighbor the largest available label 5. In the next step label 3 has at each of the alternatives two acceptable sites, leaving the last unlabeled vertex to be 4. So we end in this case with four possible labelings. With completed assignment of the labels one can construct the corresponding adjacency matrices. We immediately obtain a single matrix

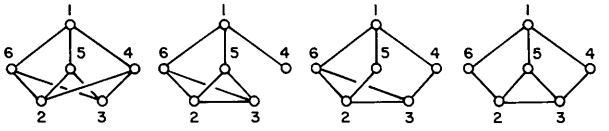
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

which describes all four distinctive labelings. As is known, this only means that the particular graph has equivalent vertices, but since here we are not interested in the symmetry properties of the graph we may adopt any of the resulting labelings. We have completed our search and have obtained the unique adjacency matrix for the first of the three graphs to be tested for isomorphism. The process is now repeated for the second graph, and finally for the third graph, and it is clear that in this case we will again obtain the same matrix shown above. One therefore concludes that the graphs under examination are isomorphic; hence, the last two can be eliminated from further examination since they are duplicates and correspond to larger binary codes. When considering relatively simple graphs it may be more expedient to use pictorial representations of partial graphs in recognizing impossible graphs (*i.e.* those violating the valency requirements) and duplicates – which is not an analytical procedure. However, it is important also to realize that pictorial representations are not essential, and a purely analytical formalism is available.

We now proceed to search for an acceptable third row of the unique matrices. The constraints (symmetric matrix and zero diagonal elements) limit the number of possibilities to only four, one from the first case and three from the second case:

000111	000111	000111	000111
000111	001011	001011	001011
000111	010011	010101	010110.

The corresponding fragments of graphs associated with the above partial adjacent matrices are:



The first (partial) matrix completely determines the first graph, which is known as $K_{3,3}$, a bipartite graph of Kuratowski (1930) important in characterization of non-planar graphs. That the graph construction has been completed for this case can be deduced analytically from a count of the number of edges so far constructed: this is given by the number of non-zero entries above the main diagonal and should be nine, as it is.* The second matrix corresponds to a graph which cannot satisfy the valency requirement, as the pending vertex (4) cannot satisfy its required valency of three. That the second graph is unacceptable can also be shown by an examination of the partial adjacency matrix:

0	0	0	1	1	1
0	0	1	0	1	1
0	1	0	0	1	1
1	0	0	0	.	.
1	1	1	.	0	.
1	1	1	.	.	0.

The unspecified sites in the matrix have to satisfy the condition that the sum of entries in any row or column is three (as well as that the number of entries above the diagonal is nine), but this is not possible. Finally, the last two partial graphs are isomorphic; hence, only the first has to be considered. It is easy to see that in the next step (fourth row) vertices 4 and 5 must be linked as the only possibility consistent with the requirements of the problem. Hence, in this case of six vertices we have obtained as solutions only two distinct graphs (Fig. 3):

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

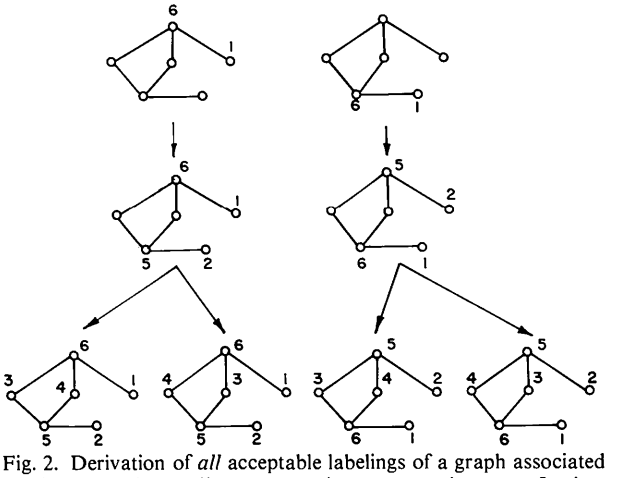


Fig. 2. Derivation of *all* acceptable labelings of a graph associated with the unique adjacency matrix as a step in a test for isomorphism. A multiple labeling only indicates the presence of equivalent vertices, the resultant matrix serves as a comparison with other matrices so derived.

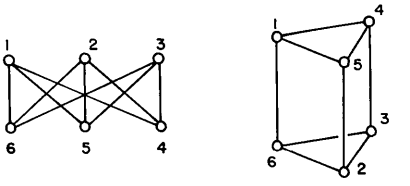


Fig. 3. The only trivalent graphs with $n = 6$ vertices.

* This is obvious: the number of entries in a particular row points to the valency of the corresponding vertex, each entry appears twice (as the adjacency matrix is symmetric) and correspond to a bond (edge). For a formalistic exposition see Theorem 2.1 in Harary (1968).

The first is the non-planar graph of Kuratowski, the second corresponds to a trigonal prism, *i.e.* a polyhedron. For a trivalent graph to correspond to a polyhedron, it must be planar and three-connected, *i.e.* it will not split into isolated fragments unless at least three vertices are excised (Steinitz, 1922).

Graphs with $n = 8$ vertices

By increasing the number of vertices, the number of combinatorial possibilities to be examined increases, as does the number of acceptable solutions. In Fig. 4 and Fig. 5 we summarize the searching process for trivalent graphs with eight vertices. At each step, which is characterized by the number of rows of the partial matrices under construction, we have eliminated cases which are isomorphic to already included partial graphs. With two rows this gives three non-isomorphic graphs as depicted in Fig. 4 at the top. Here we have already eliminated from considerations matrices with the initial rows:

0 0 0 0 0 1 1 1
0 0 1 1 1 0 0 0 .

Matrices with such a beginning would, as one can verify by a trial, lead only to duplicates. They can be *a priori* eliminated as not satisfying the fundamental requirement of corresponding to the smallest possible binary code. Namely, here the label 2 has not been used for the available site adjacent to an already assigned label (6, 7, or 8) which would reduce the number of new labels to be introduced at this stage to two. In other words, we can introduce the concept of an *effective* valency, which is determined by the number of new labels required in a partially assigned graph, and say that the above choice of assignments is unacceptable since vertices of lower effective valency available have been ignored. In fact, the existing partial assignment *dictates* that the label 2 be adjacent to label 10, as the optimum. Any other choice, like the above, or

0 0 0 0 0 1 1 1
0 0 0 1 1 0 1 0

overlooks the availability of a site that would produce a smaller binary code.

The three distinct partial graphs of the top of Fig. 4 now need to be further examined. By adding the third row we were able to construct seven partial graphs: three originate from the first graph, four from the second graph, and none from the third graph, as can be easily detected by comparing the first two rows of the new and old partial graphs. The third graph only produces duplicates which are eliminated once recognized as such. The efficiency of the procedure would be further enhanced if one could foresee such an outcome at the previous stage, but at the moment it is not clear how to recognize the situation, and in order not to omit some acceptable structures, it is imperative to avoid shortcuts that are not fully understood. The first of the seven partial graphs is unacceptable, since it has already produced an isolated fragment (all the vertices present have their valencies saturated) and

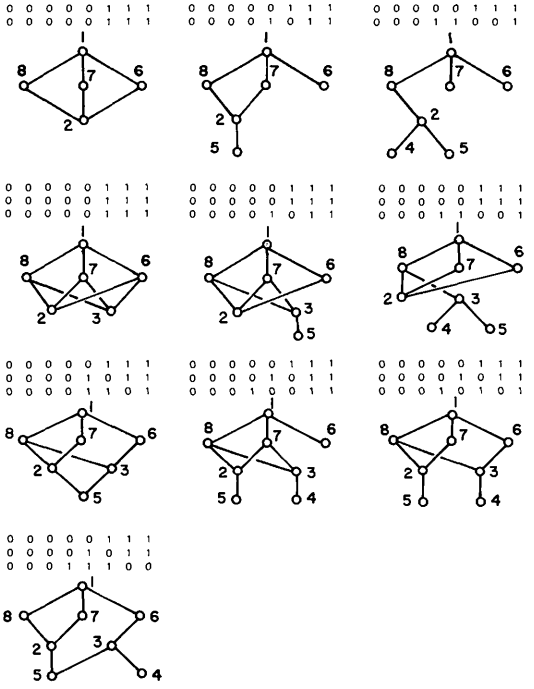


Fig. 4. The initial steps in a search for all trivalent graphs with $n = 8$ vertices.

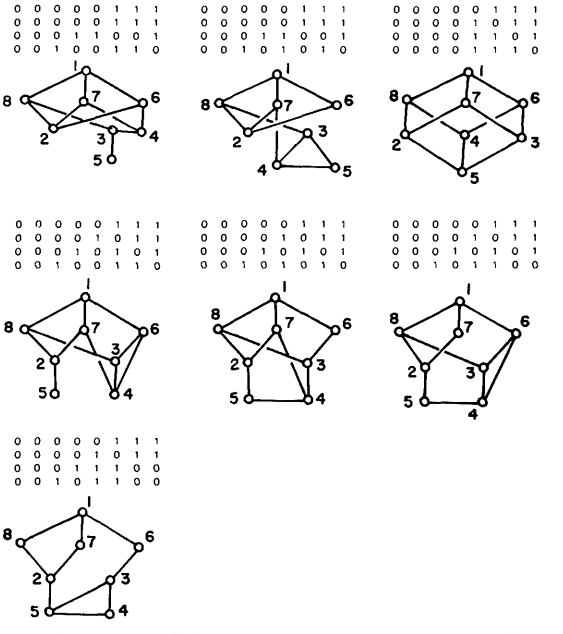


Fig. 5. A later stage of the search for the trivalent graphs with $n = 8$ vertices.

would eventually result in a disconnected graph (providing that in the continuation of the construction valency conditions for the remaining vertices could be met – which is not the case here). Similarly, the second of the seven partial graphs is not productive as a closer look reveals that the fourth vertex (to be added) cannot saturate its valency, the same being also the case with the fifth partial graph. By considering acceptable forms for the fourth row, seven new partial graphs result from the four partial graphs that yielded an acceptable alternative. These are depicted in Fig. 5, and their relation to graphs of the previous step can easily be detected by comparing the first three rows. The graph of the cube is fully completed at this step, *i.e.* all vertices and all edges have been found. All other cases have also been fully determined, but one of edges still needs to be inserted. This is evident from the pictorial forms, but can also easily be detected by summing the number of non-zero entries above the main diagonal, which should in this case be 12 if the construction is complete. In the process, possibilities that result in duplicates again have been abrogated. We left out also some permutations that may seem to have been worth consideration. For instance, after the first two partial graphs of Fig. 5 one may expect the partial matrix:

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0. \end{matrix}$$

This possibility has not been considered here because the smallest label under examination (4) has not used the largest available neighbor – hence, there is another labeling which can produce a smaller binary code (which is shown as the second in Fig. 5).

Finally, when the missing (predetermined) edge is inserted we obtain, in all, five distinctive trivalent graphs with eight vertices. Except for the first and the fourth partial graphs of Fig. 5, which have pendant vertices and therefore cannot satisfy the valency condition, all the other cases result in acceptable solutions. The resulting graphs are presented in Fig. 6 together with the associated labeling. Examination of the corresponding adjacency matrices would reveal that only the graph of the cube is bipartite. Another interesting consequence of the unique matrices is that they allow a natural (lexicographical) ordering of the derived structures. Such an order is uniquely derived by sequencing graphs by the increasing magnitude of the derived binary code. Hence, the position of a graph in the sequence is not arbitrary, and the *ordinal numbers may be viewed as structure parameters.* All regular graphs then can be represented by a symbol (v^n, k) where v is the valency, n the number of vertices, and k the position in the sequence of unique binary codes.

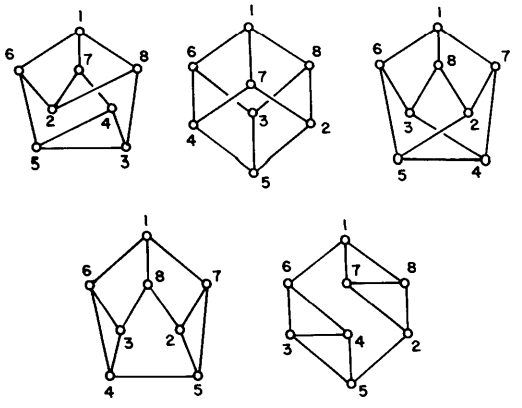


Fig. 6. Acceptable solutions (trivalent graphs with $n = 8$ vertices) with labels defining the unique matrices.

Graphs with $n = 10$ vertices

As another example of the procedure, we report the results for regular trivalent graphs with $n = 10$ vertices. These are still relatively small graphs, and most of the time the required test for isomorphism can be replaced by visual inspection. Recognition of equivalent vertices, and edges, is of help, but symmetry properties of graphs may not be apparent (Randić, 1976; Dunitz & Prelog, 1968). In Fig. 7 we have summarized the results of the search for trivalent graphs with ten vertices. In all there are 19 distinctive graphs (shown together with the corresponding labeling of vertices) which are already ordered according to the increase of the magnitude of the unique binary code. It is interesting that in a number of instances graphs with a common fragment are in proximity in the sequence, as are the pairs 2,3; 7,8; and 15,16. Some of the properties of the graphs of Fig. 7 are summarized in Table 1.

Discussion

The number of trivalent graphs has increased from 5 with $n = 8$ vertices to 19 with $n = 10$ vertices, which indicates a rapid growth with the increase of the number of vertices involved. A closer comparison of the graphs with $n = 8$ and those with $n = 10$ vertices shows

Table 1. *Properties of the graphs of Fig.* 7

Planar:	(4), (6), (7), (8), (11), (14), (15), (16)
Non-planar:	(1), (2), (3), (5), (9), (10), (12), (13), (17), (18), (19)
Bipartite:	(2), (5)
Polyhedra:	(6), (7), (8), (14)
1-connected:	(4)
2-connected:	(1), (11), (15), (16)
3-connected:	(2), (3), (6), (8), (10), (13), (14), (18), (19)
Other:	(5), (7), (9), (12), (17)

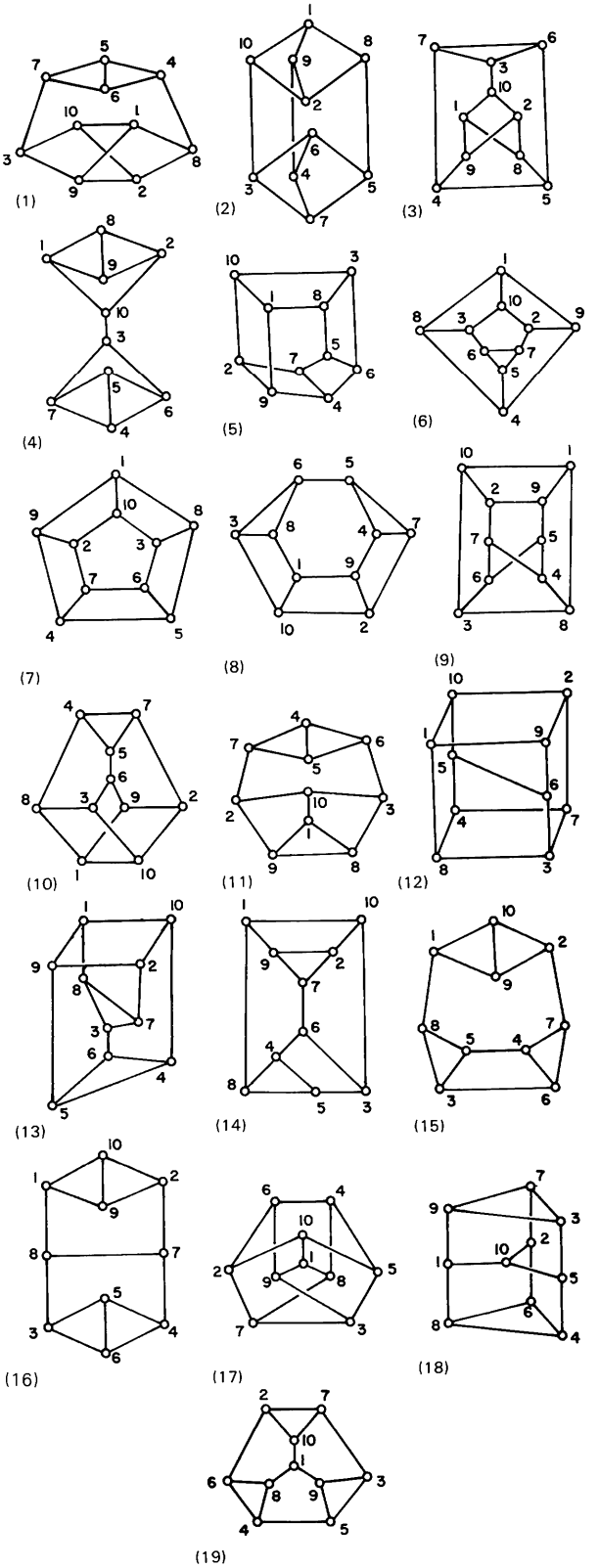


Fig. 7. Acceptable solutions (trivalent graphs with $n = 10$ vertices) with the labels defining the unique matrices and *ordered* by the magnitude of the corresponding binary code.

that a number of the latter can be considered as derived from the former.

By examining the individual graphs in Fig. 7 we observe that some graphs are 1-connected, some are 2-connected and some are 3-connected. A graph is said to be n-connected if removal of n vertices with their incident edges produces a disconnected residual. We may now consider as characteristic fragments of a graph those fragments that can be isolated by such removal of a vertex or vertices with incident edges. Since the fragments themselves may be viewed as smaller graphs (with variable valency), we see that a sizable portion of trivalent graphs can be generated in a relatively straightforward way from smaller graphs. This is important to recognize, since the effort for construction of larger graphs which are regular and of the degree 3 may then be concentrated on construction of those graphs which cannot be generated in this way, *i.e.* graphs that require more than three cuts to be disconnected. We will refer to these as *cages* or caged structures, which is close to an intuitive use of the term in organic chemistry.* The smallest such trivalent graph is the cube itself, which appears as a cage. We have redrawn the five cases in Fig. 8 to show that they all can be viewed as consisting of two five-membered rings which are interconnected in different modes. It

* The term is used in the mathematical literature in more restrictive sense, cf. Harary (1968, p. 174).

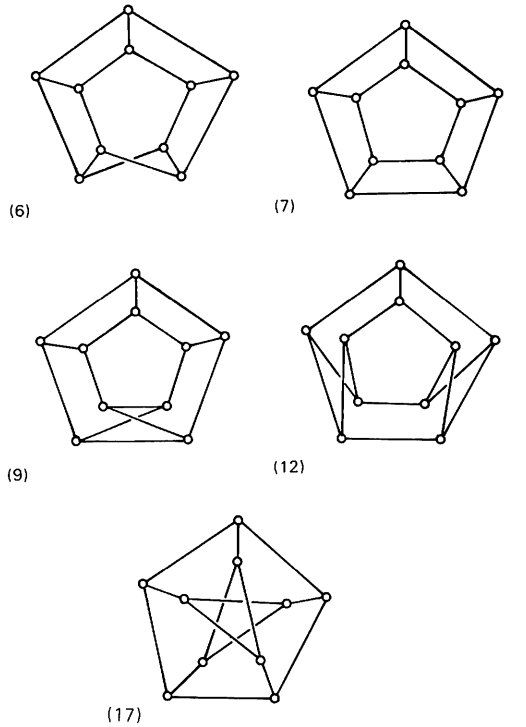


Fig. 8. The five trivalent graphs with $n = 10$ vertices which cannot be fragmented by cutting three or fewer edges.

remains to be seen if similar regularities will also hold for graphs with $n = 12$ and more vertices.

The purpose of this work has been to demonstrate an analytical route to generation of graphs of prescribed form. In the series of papers that will follow the subject will be elaborated and applied to problems of interest in chemistry, such as construction of all molecular skeletons of prescribed valency distribution, construction of all polyhedra of a given size, construction of cluster forms of interest in chemistry of heavy elements, and construction of two-dimensional and three-dimensional connected networks of interest in crystal chemistry. The application can be extended to structures having multiple bonds and loops, as well as to directed graphs when appropriate modifications are introduced which take into account the new constraints on the adjacency matrices. Clearly, some ramifications might be of more interest for their mathematical or chemical, rather than their structural, aspect (for example, enumerating polyhedra or various isomers); as these arise they will be briefly mentioned, and the interested reader will be referred to more complete accounts of such studies presented elsewhere.

The author wishes to thank Dr D. H. Rouvray (Mülheim, Germany) for comments which led to an improvement of the presentation of the material. This work was supported by the US Department of Energy, Division of Basic Energy Sciences.

References

BALABAN, A. T. (1966). *Rev. Roum. Chim.* **11**, 1097–1116.
BALABAN, A. T. (1967). *Rev. Roum. Chim.* No. 1 (last, unnumbered, page).
BALABAN, A. T. (1970). *Rev. Roum. Chim.* **15**, 463–485.
BALABAN, A. T. (1972). *Rev. Roum. Chim.* **17**, 865–881; 883–896.
BALABAN, A. T. (1973). *Rev. Roum. Chim.* **18**, 635–653.
BALABAN, A. T. (1974). *Rev. Roum. Chim.* **19**, 1185–1196; 1611–1619.
BRÜCKNER, M. (1900). *Vielecke and Vielfläche.* Leipzig: Teubner.
DUNITZ, J. D. & PRELOG, V. (1968). *Angew. Chem.* **80**, 700.
GRACE, D. W. (1965). *Computer Search for Non-isomorphic Convex Polyhedra.* Stanford Computational Center Technical Report No. CS15.
HARARY, F. (1968). *Graph Theory.* Reading, Mass.: Addison-Wesley.
KURATOWSKI, K. (1930). *Fundam. Math.* **15**, 271–283.
LEDERBERG, J. (1965). *A System for Computer Construction, Enumeration and Notation of Organic Molecules as Tree Structures and Cyclic Graphs.* NASA Report CR-68898.
LEDERBERG, J. (1966). *Systematics of Organic Molecules, Graph Topology and Hamilton Circuits.* NASA Report CR-68899.
MACKAY, A. L. (1975). *J. Chem.* Phys. **62**, 308–309.
RANDIĆ, M. (1974). *J. Chem. Phys.* **60**, 3920–3928.
RANDIĆ, M. (1976). *Chem. Phys. Lett.* **42**, 283–287.
RANDIĆ, M. (1977). *J. Chem. Inf. Comput. Sci.* **17**, 171–180.
STEINITZ, E. (1922). In *Encyklopädie der Mathematischen Wissenschaften*, IIIAB12, p. 1.

Acta Cryst. (1978). A**34**, 282–284

Stable Calculation of Coordinates from Distance Information

BY GORDON M. CRIPPEN AND TIMOTHY F. HAVEL

Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco 94143, *USA*

(*Received* 12 *August* 1977; *accepted* 18 *October* 1977)

A new method is described for the calculation of Cartesian coordinates for n points given the $n \times n$ matrix of interpoint distances. The algorithm is faster than some earlier methods, and it is remarkably stable with respect to both numerical roundoff errors and errors in the given distance matrix. The resultant coordinates have their origin near the center of mass and axes approximately along the three principal rotational axes. The calculation is described of distances to the center of mass directly from the distance matrix. Results of computer trials of the algorithm are given.

Introduction

Our recent work (Crippen, 1977*a*,*b*; Kuntz, Crippen & Kollman, 1977) on the calculation of the conformation of proteins by the 'distance geometry' approach has as an important step the computation of atomic coordinates given a trial matrix of interatomic distances. The $n \times n$ trial matrix **D** is chosen to be elementwise bounded by a matrix of upper bound distances **U** and one of lower bound distances **L**, but even so, it is usually