Enumeration and Generation of Three-Center Valence Structures

Steven Leibowitz,[1a] Irving R. Epstein,*[1a] and Daniel J. Kleitman[1b]

Contribution from the Department of Chemistry, Brandeis University, Waltham, Massachusetts 02154, and the Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. Received November 12, 1973

Abstract: The problem of enumerating and generating valence structures for molecules containing both two-center and three-center bonds is discussed. Two formal solutions in terms of an integer programming problem and a polynomial representation are suggested. A construction algorithm and a chemically suggestive variant of it are developed and discussed.

In electron-deficient molecules, particularly in boron hydrides and carboranes, topological methods which utilize the three-center bond concept[2] have given useful descriptions of both structure[3,4] and reactivity.[5] An important element of such methods is the enumeration and/or generation of all allowed three-center valence structures for a given configuration of atomic nuclei.

The problem of counting Kekulé structures in conjugated hydrocarbons was considered by Cvetković, Gutman, and Trinajstić,[6] who obtained an expression for the number of such structures as the permanent of an adjacency matrix. Hosoya[7] has given a polynomial representation for the numbers of Kekulé and Dewar structures of even alternant hydrocarbons.

The problem which we consider here[8] is: can we obtain analogous, relatively simple procedures for enumerating valence structures for boron hydrides, in which there are three-center as well as two-center bonds? A computer program for the enumeration of such structures has been developed by Epstein.[4] However, the approach employed was both enumerative and exhaustive, in that all possible combinations of two- and three-center bonds must be generated and evaluated. We seek here approaches that are more efficient timewise and hence render isomer enumeration practicable and convenient for quite large molecules even without the aid of a computer.

In the next section we describe the three-center valence structure enumeration problem. We then obtain results of two kinds. First we exhibit two alternative formal representations of the enumeration problem as a covering problem. Then we describe a constructive algorithm for performing the enumeration that renders the problem accessible to hand computation even for molecules as large as $B_{16}H_{20}$, for which the computer

program required 2.5 hr of time on the IBM 360/65 in order to select the 218 acceptable structures from over 2.5 million possibilities.[9] An example of the use of the algorithm is given. Finally, another representation of the problem, for which an analogous algorithm may be deduced, is suggested and some properties of possible chemical significance are pointed out.

I. The Boron Hydride Structure Problem

A given boron hydride molecule may be described by a graph,[10] the atomic configuration graph (ACG), whose vertices represent boron nuclei and each of whose arcs or edges connects a pair of nuclei that may participate in a common bond. Examples of some ACG's are given in Figure 1. The complexity of the situation arises from the fact that there are four types of bonds, only two of which are uniquely determined by the arrangement of the nuclei. The four kinds of bonds follow.

1. Boron–hydrogen (B–H terminal) bonds. Most boron atoms are bound to a single terminal hydrogen, though some are bound to zero or two. These bonds are considered to be determined by the given atomic positions.

2. Boron–hydrogen–boron (B–H–B bridge) bonds. These bonds are also determined by the given molecular structure or ACG. They are represented by arcs (curved lines) in the ACG's and valence structure graphs.

3. Boron–boron (B–B two-center) bonds. Such bonds may join any pair of vertices (boron atoms) that are connected by an edge (straight line) in the ACG.

4. Boron–boron–boron (B–B–B central three-center[11]) bonds. These bonds may join any three vertices that are connected by three arcs or edges in the ACG. Such a set of three vertices will be referred to as a *triangle*.

A valence structure for a boron hydride consists of a

(1) (a) Brandeis University; (b) Massachusetts Institute of Technology.

(2) (a) H. C. Longuet-Higgins, *J. Chim. Phys. Physicochim. Biol.*, **46**, 275 (1949); W. H. Eberhardt, B. L. Crawford, Jr., and W. N. Lipscomb, *J. Chem. Phys.*, **22**, 989 (1954); (b) for a recent review, see W. N. Lipscomb, *Accounts Chem. Res.*, **8**, 257 (1973).

(3) R. E. Dickerson and W. N. Lipscomb, *J. Chem. Phys.*, **27**, 212 (1957).

(4) I. R. Epstein and W. N. Lipscomb, *Inorg. Chem.*, **10**, 1921 (1971).

(5) I. R. Epstein, *Inorg. Chem.*, **12**, 709 (1973).

(6) D. Cvetković, I. Gutman, and N. Trinajstić, *Chem. Phys. Lett.*, **16**, 614 (1972).

(7) H. Hosoya, *Theor. Chim. Acta*, **25**, 215 (1972).

(8) For recent discussions of related problems in isomer enumeration, see W. Klemperer, *J. Amer. Chem. Soc.*, **94**, 6940 (1972); D. H. McDaniel, *Inorg. Chem.*, **11**, 2678 (1972).

(9) W. N. Lipscomb, *Pure Appl. Chem.*, **29**, 493 (1972).

(10) Strictly speaking, one is dealing with a *hypergraph*, rather than a graph in the usual sense, since there are two distinct binary relations among the borons: direct bonding and bonding through a bridge hydrogen.

(11) In this discussion, we shall omit consideration of the open three-center B–B–B bond which joins sets of three vertices connected by two edges in the ACG. Such bonds appear to be necessary to describe carborane valence structures (I. R. Epstein, D. S. Marynick, and W. N. Lipscomb, *J. Amer. Chem. Soc.*, **95**, 1760 (1973)) but do not appear to play a significant role in the boron hydrides.

set of two-center and three-center bonds, or alternatively of edges and triangles, together with the given B–H terminal and B–H–B bridge bonds. An *allowed* valence structure must satisfy the following "conservation conditions."

1. Each edge in the ACG must be associated with a two-center or three-center bond. It may be associated with more than one three-center bond, but not with both a two-center and a three-center bond[12] (conservation of topology).

2. Each boron vertex must take part in a total of four bonds of all four kinds (conservation of orbitals or valency).

3. The number of framework (B–B or B–B–B) bonds of each type is determined by the number of electrons in the molecule or ion (conservation of electrons or charge). In the *styx* notation of Lipscomb,[13] the boron hydride B_pH_{p+q+c} with charge c, s B–H–B bonds, and x BH_2 groups must possess precisely $t = p + c - s$ B–B–B and $y = s - 2c - q/2$ B–B bonds.

To summarize, then, an allowed valence structure for a given boron hydride consists of a set of t triangles and y edges which cover each of the edges of the ACG at least once and such that (a) no edge is covered by both a triangle and an edge and (b) each vertex v_j is covered $4 - X_j$ times, where X_j is the number of B–H and B–H–B bonds associated with v_j. Typically, most of the framework bonds will be of the three-center type.

II. Alternative Formulations

There are a number of ways to formulate the problem described in the previous section. In the following section, we shall show that there exists an approach that is sufficiently suited to the nature of the problem's constraints as to render computation relatively easy. We begin here with two alternative descriptions of the problem which lead to feasible, if not necessarily optimal, methods of attack.

First, we note that the problem may be formulated as an integer program.[14] Let us denote the vertices of the ACG by v_i ($1 \le i \le p$), the edges by a_{ij}, and the allowed triangles by t_{ijk} where the subscripts on the a's and the t's refer to the vertices involved in the edge or triangle considered.

With each edge a_{ij} or each triangle t_{ijk} we associate a variable y_{ij} or y_{ijk}, respectively. For all v_i, a_{ij}, and t_{ijk} in the ACG we demand that the variables satisfy the following constraints.

$$0 \le y_{ij} \le 1$$

$$0 \le y_{ijk} \le 1$$

$$y_{ij} = y_{ji} \quad (y_{ij} \text{ is symmetric})$$

$$y_{ijk} = y_{jki} = y_{kij} = y_{jik} = y_{kji} = y_{ikj}$$

(y_{ijk} is unchanged by cyclic permutation of its indices)

$$y_{ij} + \sum_k y_{ijk} \ge 1$$

$$y_{ij} + y_{ijk} \le 1$$

(12) In some molecules, particularly carboranes (see ref 11), one may wish to relax this last restriction and consider so-called "π-donation" structures.

(13) W. N. Lipscomb, "Boron Hydrides," W. A. Benjamin, New York, N. Y., 1963, p 45.

(14) H. Greenberg, "Integer Programming," Academic Press, New York, N. Y., 1971.
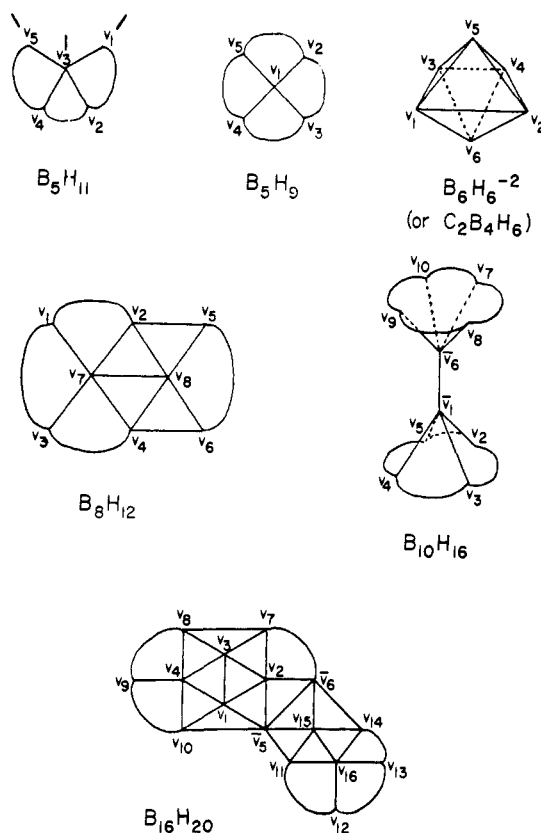

Figure 1. Some atomic configuration graphs for boron hydrides. v_n represents a boron with a single B–H bond, $\bar{v}_n$ a boron with no B–H bonds, and v_{n-} a boron with two B–H bonds. B–H–B bridges are denoted by arcs v_nv_m.

$$\sum_j \left(y_{ij} + \sum_{k>j} y_{ijk} \right) = 4 - X_i$$

$$\sum_{i<j} y_{ij} = s - 2c - q/2$$

$$\sum_{i<j<k} y_{ijk} = p + c - s$$

The number of lattice points (sets of integral values for the $\{y_{ij}\}$ and $\{y_{ijk}\}$) which satisfy all the constraints is the desired answer, the number of allowed valence structures. We may interpret a_{ij}'s or t_{ijk}'s whose variables take the value 1 at a particular lattice point as being those bonds present in the corresponding valence structure.

Since all the above constraints are linear, the problem is a kind of integer program. In the usual integer programming problem, one has an objective function, also linear in the y's that one seeks to maximize. In the present problem, one can obtain solutions (lattice points) by introducing an arbitrary objective function, solving the optimization problem for that function, introducing a further constraint to eliminate the solution just found, and iterating the procedure.[15]

There has been considerable progress in recent years in the development of general integer programming methods,[14] and the problem considered here is within the range that can be solved with reasonable speed on modern computers. In general, one might expect to be able to find all solutions (valence structures) more easily by such methods than by exhaustive enumeration.

(15) Of course, care must be taken to ensure that the new constraint introduced eliminates *only* the solution just found and no others. In practice this poses no difficulty.

However, the most efficient ways to solve integer programming problems are those that make maximal use of the inherent structure of the problem. In other words, any use we can make of chemical intuition should make the general approach outlined above (or even an enumerative approach) far more effective.

A second method of attack is to seek a mathematical formula for the number of valence structures of a given ACG. Although the practicality of such an approach is questionable, we indicate here, for the sake of completeness, how such a formula may be derived.

We first define variables Z_i for each v_i and Z_{ij} for each edge or arc a_{ij}. We also define a polynomial F ($\{Z_i\} \{Z_{ij}\}$, φ_1, φ_2) by the expression

$$F = \prod_{\substack{t_{ijk} \\ i<j<k}}(1 + \varphi_1 Z_i Z_j Z_k Z_{ij} Z_{ik} Z_{jk}) \prod_{\substack{a_{ij} \\ i<j}}(1 + \varphi_2 Z_i Z_j Z_{ij})$$

Each monomial term in F corresponds to a particular choice of edges and triangles. Those of interest here are terms having degree $p + c - s$ in φ_1 and $s - 2c - q/2$ in φ_2, degree $4 - X_j$ in each Z_j, and degree at least 1 in each Z_{ij}. These coefficients may be extracted by performing suitable integrals over F, which can be made to give rise to an explicit formula for the number of valence structures in question.

III. The Construction Algorithm

The method described in this section represents an attempt at systematizing the intuitive "trial and error" approach of boron chemists who have been constructing three-center valence structures for many years. It may be termed a systematic branching procedure for the generation of valence structures, in the sense that the set of all such structures may be branched or divided into a subset of structures possessing a particular edge or triangle and a complementary subset in which the edge or triangle is absent. Within a branch, i.e., among solutions in either of the two subsets defined by the particular branching, it is often possible to reduce the problem in size and/or to deduce the presence or absence of other edges or triangles in the solutions which belong to that branch.

By a *systematic* branching procedure we mean a sequence of branchings and reducing steps that ultimately lead to all possible solutions. To be systematic the procedure should contain both indications as to which edge or triangle one is to branch on at any stage and also a simple, well-defined reduction procedure that may be applied on each branch.

Such a procedure is efficient if it leads to dead end branches only rarely. In that case most branches will actually yield solutions, so that the number of reductions required to enumerate the N solutions will be of the order of N times the number of reductions necessary to define a solution. Since this number is generally of the order of p (the number of boron atoms) and since each reduction requires at most on the order of p steps, one can perform the enumeration (if the procedure is efficient) in on the order of $p^2 N$ steps, as opposed to a number of steps exponential in p required by the complete enumeration of configurations followed by elimination of those violating constraints.[16] We find consider-

(16) Although the computer program described in ref 4 makes considerable use of the structure of the problem and hence quickly eliminates many configurations, it is still an enumerative approach and its time requirements do appear to be exponential in p.

ably better performance in practice than the estimate of $p^2 N$ above.

Our systematic procedure may be described as follows. 1. Draw the ACG for the molecule. For convenience, choose an edge a_{ij} which is a member of one and only one triangle t_{ijk}. Such edges will be referred to as external edges. Branch on this edge; that is, divide the solutions into those which contain a_{ij} and those which do not. The latter branch must therefore contain t_{ijk}. If there are no external edges choose an edge at random and branch into those structures which contain the edge and those which do not.

2. Structures containing an external edge a_{ij} may be viewed as solutions to a reduced problem in which the edge a_{ij} is omitted and the degrees of vertices v_i and v_j are reduced by one. (Initially their "degrees" are $4 - X_i$ and $4 - X_j$.) Solutions containing a triangle t_{ijk}, where the edge a_{ij} is external, may be considered as solutions to a reduced problem in which the edge a_{ij} is omitted, the edges a_{ik} and a_{jk} are converted to B–H–B arcs, and the degrees of v_i, v_j, and v_k are each reduced by 1. If a branch contains only structures lacking a triangle incident to a particular B–H–B bond, then that bond may be suppressed. When the degree of a vertex has been reduced to zero, that vertex and all edges or arcs associated with it are removed from the graph.

3. (a) If any vertex of a (reduced) graph has degree less than half the number of edges emanating from it, then the branch contains no solutions, since it will be impossible to cover all the edges.

(b) If in a (reduced) graph a vertex has degree one or two and has incident to it only one triangle and two edges, then one can reduce the graph further by observing that either the triangle (if the degree is one) or both edges (if the degree is two) must be in all associated solutions and applying the reduction procedures of step 2 to these elements.

As will be seen in the examples, the use of molecular and local symmetry can greatly increase the efficiency of these procedures. Also, the use of systematic reduction techniques suggests the utility of having precompiled sets of solutions for relatively small, frequently occurring subunits (reduced graphs).

As a first example we consider the $B_{10}H_{13}^-$ ion. The procedure is illustrated in Figure 2. By the conservation of electrons, all solutions must contain three B–B and six B–B–B bonds. We start, according to rule 1, at the external edge a_{78} and obtain, by rule 2, the reduced graphs corresponding to the a_{78} and t_{378} branches. In the a_{78} branch, rule 3b applied at vertices 7 and 8 requires the presence of t_{848} and t_{237}, which in turn prohibits inclusion of t_{489} and t_{367}, permitting a further reduction of the graph by elimination of v_7 and v_8. Rule 3b applied at v_9 now requires a_{49} and a_{910}. Since all three B–B bonds have now been accounted for, a_{410} must be covered by t_{1410} and a_{510} by t_{1510}. The resulting reduced graph requires t_{256} and t_{123} to cover the remaining edges a_{12}, a_{13}, a_{25}, and a_{26}, thus completing the structure, the unique allowed structure containing the edge a_{78}. Analysis of the t_{378} branch proceeds similarly, as shown in Figure 2. We ultimately obtain a total of 16 solutions. Note in particular that symmetric reduced graphs yield a multiplicity of solutions which depends upon the symmetry of the graph. Thus, even in a species with no molecular symmetry, use may be made
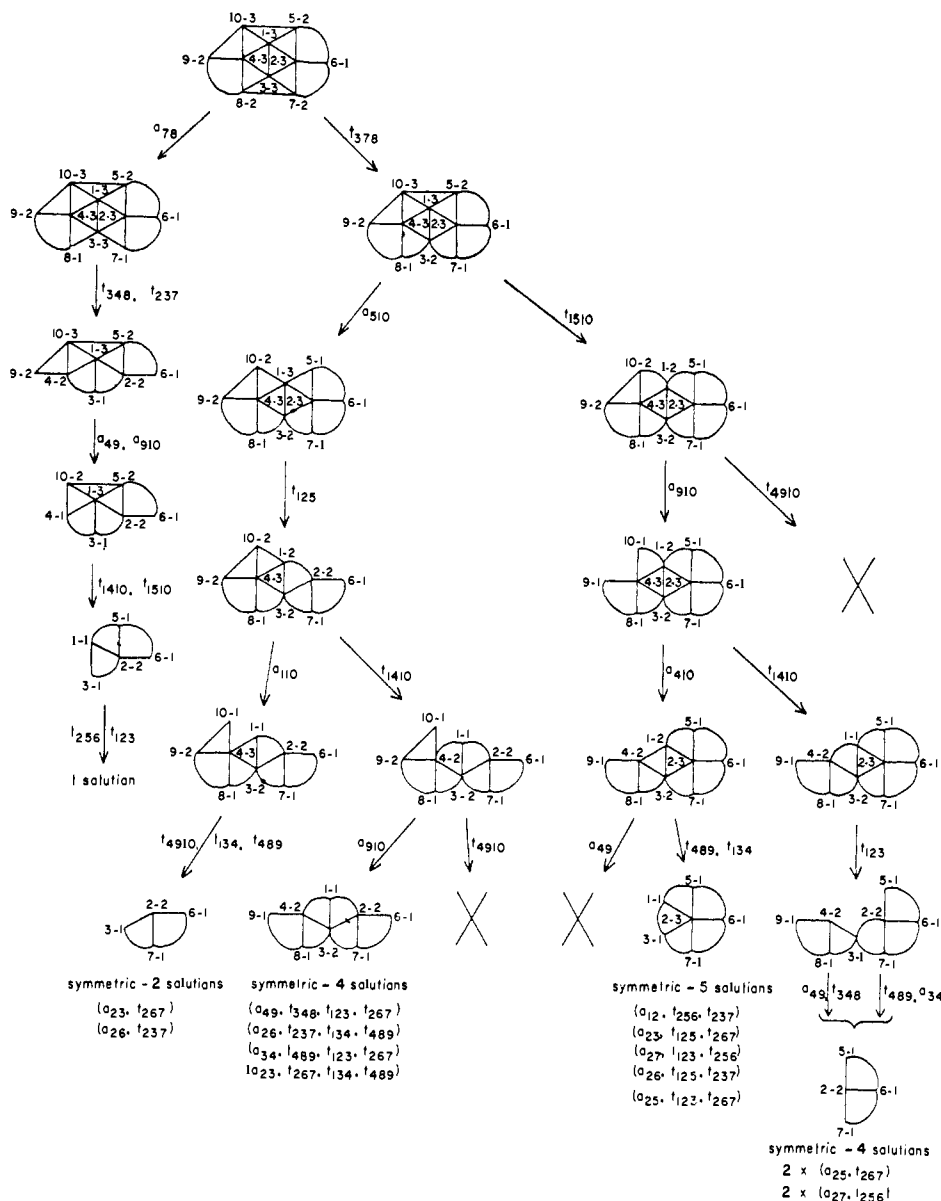
Figure 2. Branching procedure for $B_{10}H_{13}^-$ (see text). All boron atoms are attached to a single terminal hydrogen. Numbers following the vertex numbers are degrees of the vertices. Dead end branches are marked by an X.

of local symmetry. Obviously, the existence of molecular symmetry simplifies the problem further.

IV. The Derived Graph

As a variant of the construction algorithm discussed above, we indicate here how the problem may be approached by utilizing another graph, which is related to the ACG, which we shall designate as the *derived graph* of the molecule. Consider the graph derived from the ACG by taking as vertices the edges of the ACG and joining two vertices by an edge if and only if the corresponding edges in the ACG belong to a common triangle. A simple construction of the derived graph is obtained from the ACG by placing dots (vertices) at the midpoints of all edges, connecting all pairs of points found in a common triangle, and erasing the original ACG.[17] Figure 3 contains the derived graphs for the molecules whose ACG's are given in Figure 1.

It is clear that, since permutations which interchange

(17) A sheet of tracing paper is an enormous aid to this construction.

symmetry equivalent vertices of the ACG must also take edges and arcs into their symmetry equivalents, the derived graph must have symmetry equal to or greater than the ACG. Also, triangles in the ACG correspond to triangles, edges, or vertices in the derived graph according to whether the triangle contained three, two, or one edge, respectively, in the ACG. A solution to the problem in terms of the derived graph is simply a set of vertices and triangles which conform to the conservation constraints discussed in section I. An algorithm analogous to that of the previous section may easily be derived for choosing vertices and triangles from the derived graph.

Some classes of problems become considerably easier to solve in the derived graph representation. In particular, when the number of edges in the ACG is comparable to or smaller than the number of boron atoms, the derived graph is generally simpler to work with. As an extreme example, the derived graph for B_4H_{10} consists of a single isolated vertex. Perhaps a more im-
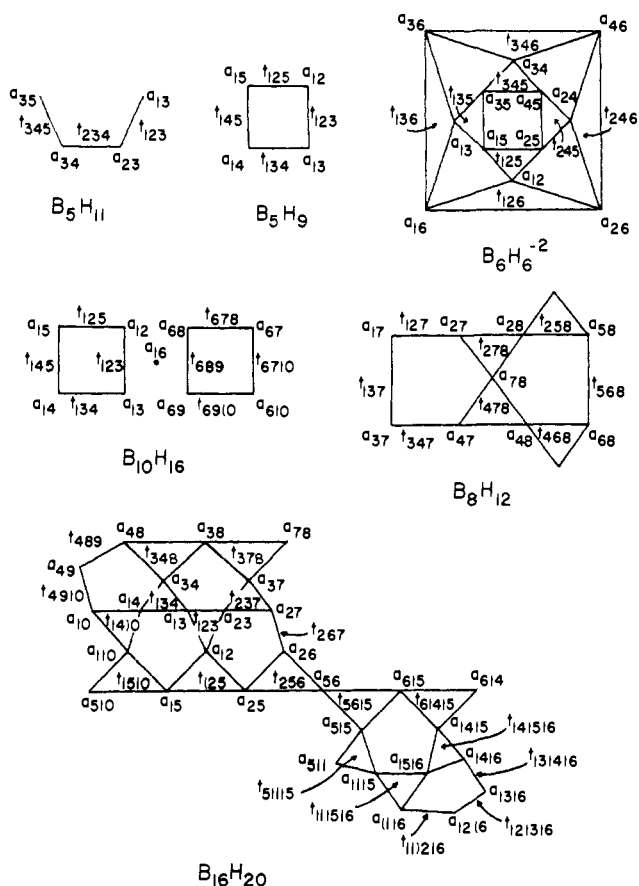
Figure 3. Derived graphs corresponding to the boron hydrides of Figure 1. Edges and triangles are labeled according to the vertex numbering scheme of the ACG's in Figure 1.

portant class of problems, particularly for the larger boron hydrides, is that in which the set of solutions may be factored into two parts, *i.e.*, those problems in which the molecule consists of two pseudoindependent subunits. Although this factorization may often be obtained from direct consideration of the ACG, it is generally more apparent in the derived graph. Two examples of this phenomenon appear in the derived graphs of $B_{10}H_{16}$ and $B_{16}H_{20}$ in Figure 3. In the former case, the fact that all valence structures may be expressed as products of two B_5H_9 structures plus a_{16} is immediately apparent from the existence of the three disjoint subgraphs. The "bottleneck" at a_{56} in the $B_{16}H_{20}$ derived graph allows it to be separated into a B_8 subunit whose derived graph is isomorphic to that of B_8H_{12} or a related ion (*e.g.*, $B_8H_{13}^-$) and a B_{10} subunit isomorphic to $B_{10}H_{14}$ or $B_{10}H_{13}^-$, Figure 2. This separation made it possible for a student to obtain by hand the 218 valence structures for $B_{16}H_{20}$ in less time than required by the computer program.[4,9]

Finally, we note that (except in the rare instances such as $B_{10}H_{16}$ when the derived graph is disjoint) a boron atom (ACG vertex) is represented in the derived graph by a set of connected vertices (the *atomic set*), each of which contains that atom in the ACG. It is interesting

to divide the boron atoms in a molecule into two categories, according to the topology of their atomic sets in the derived graph. Those whose atomic sets are closed polygons will be called *internal*, while those whose atomic sets are line segments will be called *external*. It appears that boron atoms at internal positions tend to be more negatively charged and that the most stable carboranes are those with the carbons at external positions where possible. If one divides boron hydrides according to *closo* (*e.g.*, $B_6H_6^{2-}$), nido (*e.g.*, B_5H_9), arachno (*e.g.*, B_5H_{11}) classification,[18] one finds that *closo*-boranes contain *only* internal boron atoms. Since there are many more ways of covering internal atomic sets than external ones, this observation provides a rationalization, in terms of an increased number of valence structures, for the observed stability of the *closo*-boranes and -carboranes.[19] The *nido*-boranes have both internal and external borons, though generally more of the latter, while the smaller *arachno*-boranes have even fewer internal atoms, with those containing fewer than seven borons having only external positions. Interestingly, for the larger boranes (ten or more borons) the numbers of internal positions become comparable for nido and arachno species containing the same numbers of borons. We suggest that the distinction between internal and external positions may prove a useful one in understanding and predicting the structure and reactivity of the boranes and carboranes.

V. Conclusion

We have shown that a number of alternative, but ultimately equivalent, methods may be used to solve the problem of generating and/or enumerating three-center valence structures. The most efficient methods will be those which make maximal use of chemical intuition as opposed to mathematical "brute force." While the discussion here has dealt only with a particular type of three-center valence structure, we have attempted to keep the treatment sufficiently general that extensions to other problems may be easily accomplished. For example, inclusion of open three-center bonding[11] and/or π-donation structures[12] is easily allowed for by alteration of the "topology conservation" rules of section I. Consideration of the four- and higher-center bonding which may be present in some intermetallic and other compounds is somewhat more complicated but certainly within the scope of the methods presented here.

(18) R. E. Williams, *Inorg. Chem.*, 10, 210 (1971).
(19) The hypothesis (R. B. King, *J. Amer. Chem. Soc.*, 94, 95 (1971)) that closed triangulated polyhedra offer particular stability is also supported by this notion. Note that the external–internal distinction is possible (but far less obvious) without the derived graph representation.