

CONNOTATIONS INHERENT IN THE CONCEPT OF 'STRAIGHT'  
IN GEOMETRICAL AND MOLECULAR MODELS

Seymour B. Elk  
Elk Technical Associates  
321 Harris Place  
New Milford, New Jersey 07646

(Received: January 1988)

ABSTRACT: What should the word 'straight' connote in a molecule that contains three or more atoms? From a geometrical bias, note:

(a) In a planar conjugated environment only the all-trans isomer approximates the geometrical idea expressed by the word.

(b) In a saturated bond system, 'straight chain' n-alkanes might, just as correctly, be considered to be 'circular', as to be straight. In fact, one could ask: Is the effectiveness of certain reagents due to one of the conformers of this reagent having a circular geometry?

(c) With the appropriate choice of module in  $sp^2$  and  $sp^3$  environments, the heuristic and geometric concepts of 'straight' do coincide at the modular level, even though they did not at the individual bond level.

In the study of geometry, or more precisely in the process of synthesizing a Euclidean geometry, the concept of 'figure' is predicated on connecting entities called 'points' by means of other entities called 'lines' and then connecting parts of these lines (called 'line segments') to create 'planar' figures, etc. In this process there has been created

the heuristic elements of 'connecting', 'straight', 'interior', etc., as well as a class of words that have been designated as 'primitive'<sup>1</sup>. So long as we are merely formulating the system, it is possible to effectively mask nearly all problems of consistency by making the appropriate intuitive assumptions. Subtle difficulties arise only when an attempt is made to close the system and then 'analyze' any given figure within the system.<sup>2</sup>

A related set of dilemmas arise when applying the precepts of geometry to the description of molecules<sup>3</sup>; i.e., when the inherent geometry of the model formed by designating atoms as points and chemical bonds as line segments is used as a substitute for the chemical entity. Two of the more important of these are:

- (1) What is the appropriate 'embedding space'?<sup>5</sup> (The term 'embedding space' refers here to the mathematical space in which the physical molecule may be idealized without sacrificing its important chemical properties. For example, a plane is adequate for describing benzene; whereas a three dimensional space is more appropriate in the description of cyclohexane.)

and

- (2) What meaning is to be assigned to 'homaloidal' (A word that is independent of the dimension of the embedding space and is the generic word meaning 'straight' in one dimension, 'flat' in two dimensions, etc.)?

At this point it is important to distinguish between two types of geometry: A 'discrete geometry' is one in which there are only a finite number of points; consequently, the concept of 'contiguous' (next in sequence) is definable. For such a space two points are contiguous when, assuming agreement has been reached on some unique one-to-one correlation of these points with the set of integers, the two points in question are represented by successive integers. On the other hand, a 'continuous geometry' is one in which for any two points there is always another point midway between them. Consequently, the concept of contiguous, or even of

'near', is NOT definable for any continuous variable.

It is a heuristic element of most continuous geometries that the entities designated as figures will be the same dimension as that of the embedding space; e.g., although skew quadrilaterals are well-defined in 'solid geometry', they are subordinated to the subset of edges of the associated tetrahedron, with the major geometrical description focusing on the tetrahedron.

In chemistry, on the other hand, this geometrical generalization of thinking in terms of entire spaces is a poor one. To the contrary, the geometry that is most relevant is one in which chemical action takes place primarily between pairs of atoms. Consequently, our focus shall be ONLY on the union of these linear pairs. For example, in cyclobutane, which is frequently modelled as a skew quadrilateral, the tetrahedral 'superstructure' of the skew quadrilateral is of negligible (some may say zero) significance. Similarly, for most compounds, I know of no chemical significance to an imagined pathway connecting non-adjacent carbon atoms -- even though in some conformers such a metric<sup>6</sup> distance would be less than the equilibrium value of  $1.549\text{\AA}$ <sup>8</sup> between graph theoretically adjacent atoms. This is equivalent to saying that the chemical concept usually described by delocalized bonding has been represented in the model by a geometrical one.

As a general rule, the one-dimensional model of chemical bonds is adequate for most molecules;<sup>5</sup> i.e., the model being used does not take into consideration either two-dimensional or three-dimensional modules. In this model there is no chemical relevance to either the solid cube or to the interior of the six planar faces in cubane (Figure 1); rather the set of twelve edges between pairs of pairs of adjacent carbon atoms, as well as the eight line segments between carbon and hydrogen atoms, serve as the locus of all

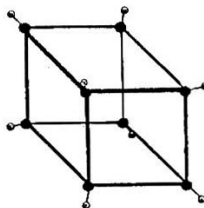


FIGURE 1

chemical properties -- with the space in-between being considered as 'empty'. More succinctly, one could say that for most compounds, such as both cyclobutane and cubane, despite the specification of a three-dimensional embedding space, a linear geometry, rather than either the planar (boundary-defined) or solid (content-defined) geometries, is the most relevant one to use in describing chemical compounds.

As an idea of major importance, observe that when we try to define the concept of 'straight' -- such as the n-alkanes are presumed to be -- an unusual description, that is contrary to our intuition, requires examination. This is a consequence of the process of forming figures by annellating line segments, in contradistinction to the more elementary process used in geometry of examining the subsets formed by the intersection of entire lines. One concomitant result is that 'straight' has lost its intuitive meaning in a system that is synthesized by annellating unstrained carbon atom bonds: The geometrical angle of connection in such a system is tetrahedral ( $109^{\circ} 28' 14''$  --  $sp^3$  orbitals) for coordination = 4 and hexagonal<sup>9</sup> ( $120^{\circ}$  --  $sp^2$  orbitals) for coordination = 3. It is only for an  $sp$  orbital that the intuitive idea of straight is in evidence.

Consider first  $sp^2$  orbitals: Note that the intuitive concept of a straight line does not exist. Instead, various methods of representing the nearest approximation to a straight line within the constraint of a hexagonal grid are presented in Figures 2, 3 and 4. In actuality, Figures 2 and 4 are merely different ways of representing the same all-trans system, while Figure 3 is an

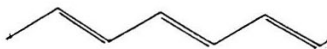


FIGURE 2



FIGURE 3



FIGURE 4

all-cis system. Furthermore, Figure 4 is shown to be identical to Figure 2 by remembering that this particular model allows for free rotation about all single bonds. Thus, by successively rotating that portion of the model to the right of each single bond through  $180^\circ$ , Figure 5 is formed.



FIGURE 5

The entire molecule may now be rotated through  $30^\circ$  to give Figure 2. At this point it should be noted that Figure 2 conforms to the conventions established by Patterson<sup>10</sup> for drawing figures and Figure 4 does not; therefore, it is the one usually chosen by chemists for representing the all-trans system.

Upon comparing an all-trans and an all-cis system, notice that there exists an important difference between the conformers formed by rotation about single bonds: Using Figure 5, note that the conformer formed by such a rotation in an all-trans system does not result in non-adjacent skeleton carbon atoms being brought significantly closer together. On the other hand, using Figures 6 and 7 it may be seen that the

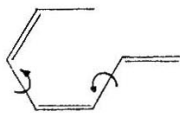


FIGURE 6

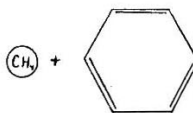


FIGURE 7

conformer formed by such a rotation in an all-cis system does. In fact, after one such rotation, a pair of non-contiguous atoms are approximately one bond length apart from each other, and after a second such rotation -- assuming that conceptually

a methane molecule could be dislodged -- a closed ring would be formed. At this point, it should be noted that, despite the logical sequence described, the compound is NOT all-cis 1,3,5-cyclohexatriene. As far as is known, such a compound does not exist. Whether such a cyclohexatriene is a 'transition state' for a tiny fraction of a second is a matter of conjecture. What is not speculation is that this particular orbital system produces a new, more stable compound (benzene) which can be considered the prototype of a different taxonomy class.<sup>11</sup> This is equivalent to suggesting that the all-cis form will tend to create benzene modules, rather than constructs of the type shown in Figure 3. Furthermore, note that an all-cis system is completely convex<sup>12</sup>, whereas each trans bond introduces concavity. Additionally, although the set of geometrical isomers of a given constitution<sup>13</sup> all have the same graph theoretical diameter<sup>15</sup>, the all-trans system has the longest metric distance, as well as intuitively having the maximum concavity<sup>17</sup>. For this reason, the all-trans construct (Figure 2) is the logical one to designate by the heuristic word 'straight'.

Turning now to  $sp^3$  orbitals: Once again the concept of straight requires interpretation. Just as hexagonal tessellation of the plane<sup>18</sup> suggested consideration of segments of the types shown in Figures 2, 3 and 4 above, let us begin this part of the study by analyzing a model of a polymantane<sup>19</sup>. See Figure 8. Here we note that an alternating sequence of line segments (two per adamantane module) fulfills the intuitive concept of 'straight'. Furthermore, the following observations from the model in Figure 8 can be made:

- (1) As in the all-trans system in a two dimensional embedding space, alternating bonds in the polymantane are parallel.
- (2) Comparable to the all-cis system in a two dimensional embedding space, a closed circuit in the polymantane is formed from six (carbon) atoms.
- (3) In each such closed circuit described in (2), the pairs of every third bond are also parallel.

Now, however, unlike the  $sp^2$  environment, all of the C-C

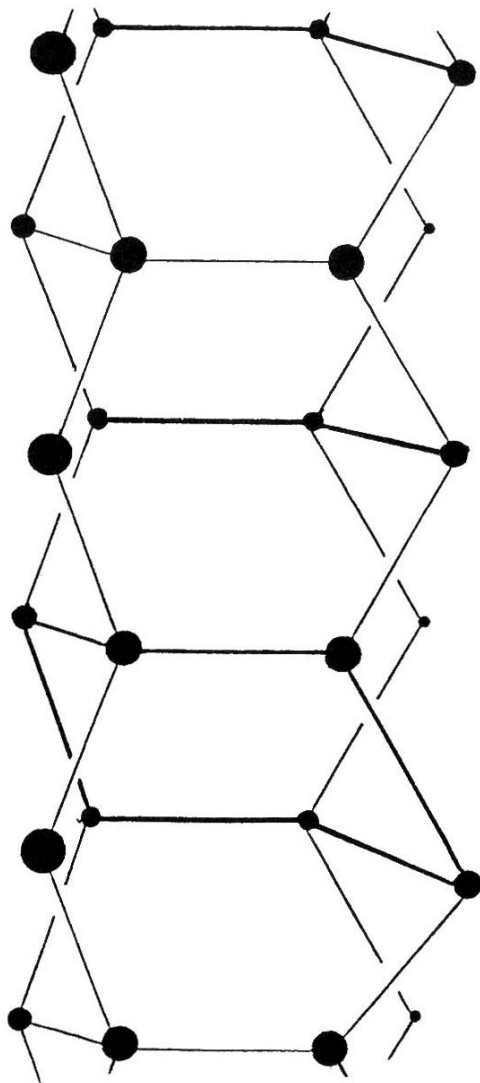


FIGURE 8

bonds in this  $sp^3$  environment are single bonds, which, were it not for the formation of rings, would offer no restriction to the free rotation about this bond. Because of this, we may envision segments of the skeleton of an n-alkane as being either 'straight' -- as exemplified by analogy to trans -- or 'circular' -- by analogy to cis. In other words, from the limited perspective of only the geometry, an active element, such as lithium, on one end of an n-alkyl chain can be involved in a reaction occurring at either one or both ends of the chain simultaneously -- depending on whether the reacting conformer is straight or circular. For example, the circular conformer of say n-butyllithium would have the requisite geometry to form a six member intra-molecular ring by acting as a 'pincer' on a single atom.

In addition to the above usage of 'straight' with respect to the bonding of atoms, there is also the need to consider 'straight' with respect to modules (composed of more than a single atom). As the simplest of these modules consider  $sp^2$  orbitals and benzene. Now the concept of straight is coincident with our intuition, producing the class of polycyclic aromatic hydrocarbons referred to as 'acenes'.<sup>20</sup>

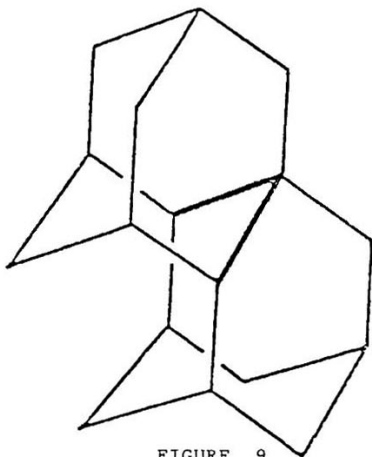


FIGURE 9

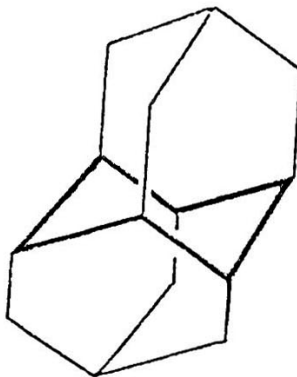


FIGURE 10



In a similar manner, for  $sp^3$  orbitals and adamantane, the concept of straight is fulfilled by both the three vertex (common pair of edges) fusion (Figure 9) and by the common face fusion (Figure 10) of adamantane modules<sup>21</sup>.

Now, so long as the (congruent) modules just described tessellate the embedding space (i.e., regular hexagons in the plane and adamantane modules in three dimensions), the intuitive concept designated by the word 'straight' is unambiguous. This, however, is not always the case. For example, because the tetrahedral angle ( $109^\circ 28' 14''$ ) is very close to the interior angle of a regular (planar) pentagon ( $108^\circ$ ), condensed five member ring systems seem to form straight molecules. In such a system, it was shown<sup>22,23</sup> that to be considered 'straight' the alternation of pentagonal rings in an up vs. down orientation was required. Nevertheless, such rings can not tessellate the plane. This is notwithstanding the fact that, using Patterson's rules for drawing these rings<sup>24</sup>, it appears that such a tessellation would be a simple matter to accomplish (Figure 11).

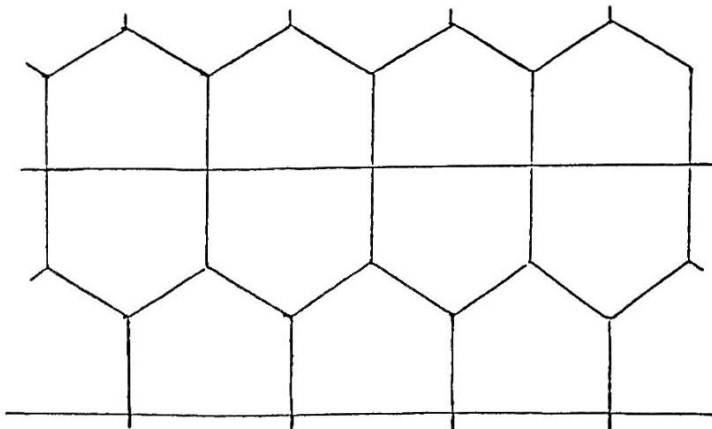
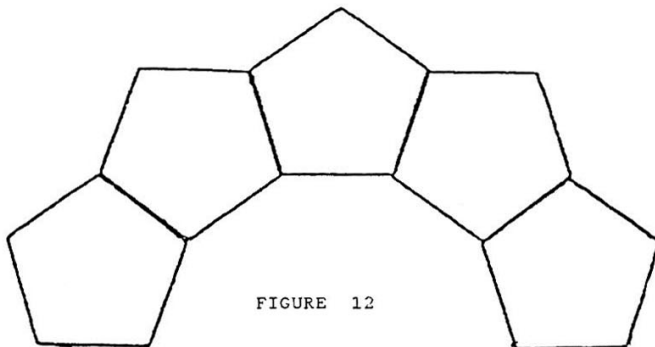
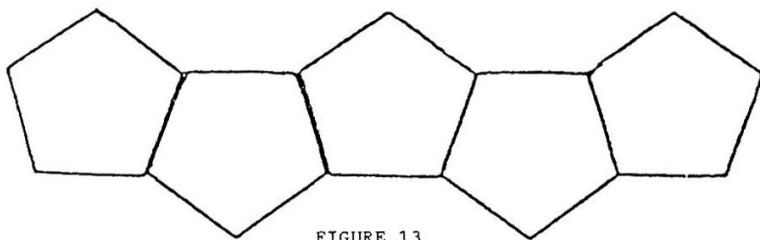


FIGURE 11

Before pursuing this idea further, note that the 'convex' orientation (Figure 12) corresponds to the 'cis' description



and produces a 'circular', rather than a 'straight', molecule; while the 'concave' orientation (although heuristically straight) creates a 'bay region' with an angle of  $144^\circ$  at the junction of two rings -- each of which contributed  $108^\circ$  (Figure 13).



Returning now to the attempt to tessellate the plane by annellating pentagonal rings in each bay region of the 'straight chain' molecule. Notice that such an annellation is possible only for non-contiguous rings. To illustrate this, let us annellate the five ring system shown in Figure 13 with a pentagon in each bay region: From Figure 14, we see

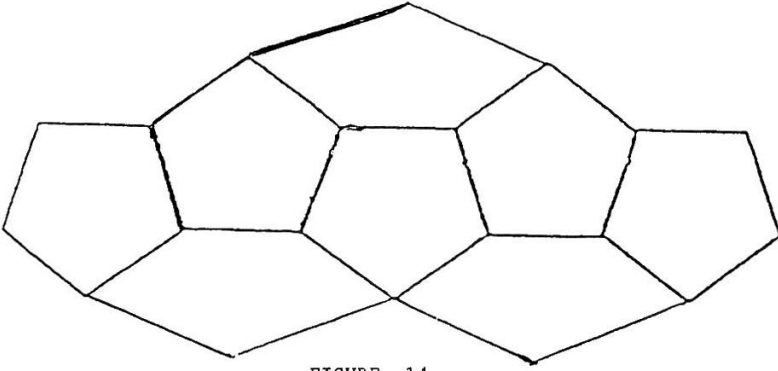


FIGURE 14

that insertion of a distorted pentagon at the 'top' of the drawing does not produce any undesirable changes in constitution. A similar insertion in the bay region at the bottom left by itself would also be acceptable; however, the insertion of two such pentagons -- which would be required if the plane was going to be tessellated -- creates an incidence = 4 vertex, whereas all previous vertices had had incidence

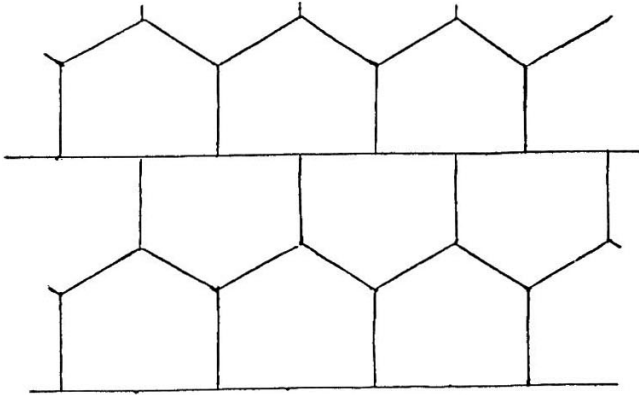


FIGURE 15

= 3. That this is not too surprising can be seen by noting that there exist carbon atoms having two different constitutions in Figure 11: incidence = 3 along the zigzag line and incidence = 4 along the heuristically straight line. The presence of an incidence = 4 carbon atom, however, is not conducive to planar structure.<sup>25</sup> More realistically, let there be slippage along the geometrical 'line' (Figure 15); thereby forming hexagons (non-regular). Now once again each vertex has incidence = 3. However, a closer examination of Figure 15 reveals that this is simply a distorted planar projection of part of a three dimensional model. By considering that a fourth bond emanates from each vertex, Figure 15 is seen to be merely one 'plane' of the polymantane shown in Figure 8. In other words, the 'simple' geometry envisioned using planar pentagons as the building module is, in actuality, simplistic. The reality is that this 'plane' formed by skew hexagonal modules in the polymantane is an example of a two-dimensional homaloid in an all  $sp^3$  environment; i.e., is 'flat'. This is equivalent to saying that just as an alternating parallel bonded union of line segments is a best approximation to the geometrical straight line in both an  $sp^2$  and  $sp^3$  environment, so the pleated union of chair-shaped hexagonal conformers of fused cyclohexane modules models the geometrical plane in the  $sp^3$  environment. For the  $sp^2$  environment, the geometrical model and the chemical representation, graphite, are coincident.

In summary, the mathematical concept of 'straight' (one dimensional and homaloidal) when applied to individual chemical bonds between atoms produces the traditional 'not straight' geometry exemplified by alternating parallel bonds. This is in contradistinction to such frequently encountered expressions as 'straight chain alkanes' -- which would be better designated as 'non-branched'.<sup>26</sup> It is only when consideration is focused on larger modules (as well as  $sp$  orbitals) that the chemical and the geometrical connotations of 'straight' are coincident.

REFERENCES AND NOTES

1. If we wish to develop any internally consistent system, some starting ('primitive') words must be accepted as understood. It is not possible to define everything. This is true since words can only be defined in terms of other words -- whose meaning is assumed to be known. Notice that a primitive word in one system may be a word which has a precise definition in another system. For example, a 'line' might be considered as primitive (as was assumed by Euclid) or else it may be defined in terms of more sophisticated concepts. One such definition -- which uses a much larger knowledge base -- is: A one-dimensional homaloidal manifold -- where the terms 'dimension', 'homaloidal' and 'manifold' are assumed to be understood by the reader and each of these terms is judged to be more primitive than 'line'. Which words you denote as 'primitive' depends on your individual bias. It is a strictly heuristic decision what is the base of knowledge and thus of reference.
2. Goedel, K., "On Formally Undecidable Propositions", Basic Books: New York, 1962
3. Turro<sup>4</sup> presents some general ideas on how geometrical and topological thinking has been physically "expressed" in chemical moieties, including the processes that go from the abstract (topological spaces) to the concrete (geometrical figures) and the influence of these and various intermediate forms on molecular structure. See Figure 7 and Sections 4.3 & 4.4 of Reference 4.
4. Turro, N. J., "Geometrisches und topologisches Denken in der Organischen Chemie", Angew.Chem. 1986, 98, 872-92.
5. Elk, S. B., "Topologically Different Models to be Used as the Basis for Ring Compound Taxonomy", J.Chem.Inf.Comput.Sci. 1985, 25, 19.
6. The term 'metric distance' (in contradistinction to 'graph theoretical distance') is the actual length between two points, measured in any convenient unit, such as Angstroms. For example, the average metric distance between two single-bonded carbon atoms is 1.54Å and between two doubly-bonded carbon atoms is 1.33Å. Graph-theoretical distance, in both cases, is the integer 1 -- no unit of dimension is used.
7. Streitwieser, A. & Heathcock, C. H., "Introduction to Organic Chemistry," 3rd Ed., Macmillan Publishing Co., Inc., New York, 1981, p.278
8. Cotton, F. A. & Frenz, B. A., "Conformers of Cyclobutane," Tetrahedron, 1974, 30, 1587-1594

9. Although the geometrical picture is that of a regular planar hexagon, chemists normally refer to this coordination as 'trigonal'.
10. Patterson, A. M., "Proposed International Rules for Numbering Organic Ring Systems", J.Am.Chem.Soc. 1925, 47, 547.
11. Elk, S. B., *ibid* #5
12. Coxeter, H. M. S., "Introduction to Geometry", John Wiley & Sons, Inc., New York, 1961, p.183.
13. "The chemical constitution of a molecule, or an ensemble of molecules, is determined by the number and kind of atoms which it contains and those pairs of neighboring atoms which are connected by covalent bonds."<sup>14</sup>
14. Schubert, W. & Ugi, I., "Constitutional Symmetry and Unique Descriptors of Molecules", J.Am.Chem.Soc. 1978, 100, 37.
15. The 'diameter' of a graph is the longest graph theoretical distance between any two vertices of that graph, where 'distance between two vertices' is defined as the shortest path; i.e., least number of edges between these vertices.<sup>16</sup> Note that this is a 'maxi-min' concept.
16. Harary, F., "Graph Theory", Addison-Wesley, Reading, Mass., 1969, p.14.
17. A quantitative measure of total concavity of a region has probably not been uniquely defined; however, in the case of molecules, the relevant ideas are the maximum number of reflex angles and the sum of the excess measure of the individual reflex angles beyond that of straight angles. Furthermore, because all of the angles involved are equal, this second attribute may be subsumed by the first.
18. Elk, S. B., "A Nomenclature for Regular Tessellations and Its Application to Polycyclic Aromatic Hydrocarbons", Math.Chem., 1980, 8, 121-158.
19. Balaban, A. T. & Schleyer, P. v. R., "Systematic Classification and Nomenclature of Diamond Hydrocarbons - I., Graph-Theoretical Enumeration of Polymantanes", Tetrahedron 1978, 34, 3599-3609.
20. Clar, E., "Polycyclic Hydrocarbons, I & II," Academic Press, New York, 1964.
21. Elk, S. B., "Effect of Taxonomy Class and Spanning Set on Identifying and Counting Rings in a Compound," J.Chem.Inf.Comput.Sci., 1985, 25, 15.

22. Elk, S. B., "An Algorithm to Identify and Count Isomeric Molecules Formed by the Linear Fusion of Cyclopentane Modules", J.Chem.Inf.Comput.Sci., 1987, 27, 67-69.
23. Elk, S. B., "Refinement of Systematic Nomenclature for Polybenzenes and Its Extension to Systems of General Arenes", Math.Chem. 1982,13, 260.
24. Patterson, A. M., *ibid* #9.
25. Greenberg, A. & Liebman, J. F., "Strained Organic Molecules," Academic Press, 1978, pp. 369-375.
26. As is common in mathematical logic, double negation can produce a widening of the field of elements over that obtained by the process of simple affirmation. Common examples of this in the ordinary use of our language include: 'not unwilling' means either 'willing' or 'indifferent', etc. Therefore, let us designate 'branched' as the negation of 'straight' in an  $sp^3$  environment. The alternates 'kinked' or 'curved' are not relevant in a topological frame of reference - only in a metric geometry. It should be noted that 'not branched' does not have the same constraining connotation as 'straight'.