The Bond Graph

By STEVEN H. BERTZ
(*Bell Laboratories, Murray Hill, New Jersey* 07974)

Summary The concept of the bond graph is introduced for the extension of 'branching' and 'topological' indices from saturated to unsaturated hydrocarbons as the key step in the construction of a general index of complexity for molecular skeletons.

GORDON and Kennedy have described how a set of graph-theoretical invariants, N_{ij}, 'the number of distinct ways in which skeleton i can be cut out of skeleton j,'† can be used as 'branching' indices to correlate the thermodynamic properties of saturated hydrocarbons.[1] One member of this set, N_{2j} (the number of ways that propane can be cut out of a skeletal[1] or hydrogen-suppressed[2] molecular graph, M'), is of special interest because it is the simplest one that reflects branching.[3] In order to construct a general index of molecular complexity useful in synthetic analysis,[4] in the correlation of properties with structure,[1,2] and in molecular evolution,[5] it is necessary to generalize this invariant to make it applicable to unsaturated systems as well.

This extension can be accomplished by considering the graph-theoretical 'first derivative,'[6] or line graph, $B(M)$, of the molecular graph M.[7] For chemical applications the term *bond graph* is a natural one,‡ as it is derived by representation of each of the lines (bonds) of M by a point and then connecting a pair of points with a line whenever the corresponding bonds are adjacent. As an immediate application, the construction of $B(M)$ from M provides a simplified basis for the derivation of the topological molecular orbitals of M.[8] Considering the definition of the bond graph and that N_{2j} can be defined as the number of pairs of adjacent bonds in a saturated hydrocarbon, N_{2j} equals the number of lines in the bond graph‡ of M', as illustrated in Figure 1 for the isomeric pentanes.

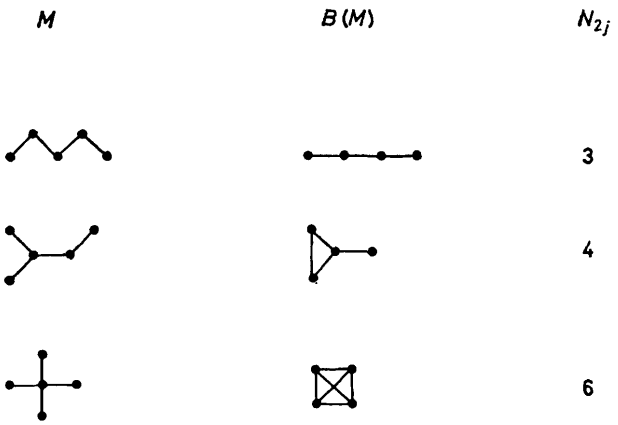

FIGURE 1. Bond graphs of the isomeric pentanes.

Previously, it was not obvious how to 'cut' propane out of multiple bonds. By defining N_{2j} as the number of lines in $B(M')$, this limitation can be overcome. Note that the alternative definition, the number of pairs of adjacent bonds, still holds because adjacency is a binary relationship.§ For example, ethylene has *one* pair of adjacent bonds, and its bond graph contains one line. Further examples are given in Figure 2. Definition of a pair of adjacent bonds as a *connection* allows the number of lines in the bond graph to be expressed as the number of connections.¶

The minimum criteria for a useful index of complexity are (i) the index must be able to treat all structural features and (ii) the index must do so in a consistent way, *i.e.* always increase as the numbers of complicating factors increase. Figure 1 shows how the number of connections increases with branching and Figure 2 shows how this number increases with substitution about a multiple bond. The examples in the Table demonstrate that it also increases with chain length (and monocyclic ring size), cyclization (the number of rings for a given number of atoms), and degree of unsaturation.

TABLE

Molecule	Connections	Merrifield–Simmons	Randić–Wilkins	Bonchev–Trinajstić	Hosoya	Randić	Rashevsky	Wiener
Ethane	0	3	1	0	2	1·00	0	1
Propane	1	5	3	2·8	3	1·41	0·92	4
n-Butane	2	8	6	8·8	5	1·91	1·00	10
n-Pentane	3	13	10	18·5	8	2·41	1·52	20
2-Methylbutane	4	14	10	15·2	7	2·27	1·92	18
2,2-Dimethylpropane	6	17	10	9·7	5	2·00	0·72	16
Cyclobutane	4	7	12	5·5	7	2·00	0	8
Cyclopentane	5	11	20	10·0	11	2·50	0	15
Cyclohexane	6	18	30	22·8	18	3·00	0	27
Bicyclobutane	8	6	19	3·9	8	1·97	1·00	7
Tetrahedrane	12	5	30	0	10	2·00	0	6
Prismane	18	13	129	14·6	32	3·00	0	21
Ethene	1	[a]	2	0	[a]	[a]	0	1
Ethyne	3	[a]	3	0	[a]	[a]	0	1
Cyclobutene	7	[a]	18	5·5	[a]	[a]	1·00	8

[a] Undefined by authors.

† In more mathematical terms 'N_{ij} is the number of subgraphs of [graph j] isomorphic with [graph i],' (see ref. 1).

‡ In addition, it avoids confusing constructions such as 'the number of lines of the line graph.'

§ Thus, the line graph of a multigraph is a simple graph (see ref. 6).

¶ This term results in a great economy of words. For saturated hydrocarbons it can also be used synonymously with 'paths of length two' (see ref. 10).

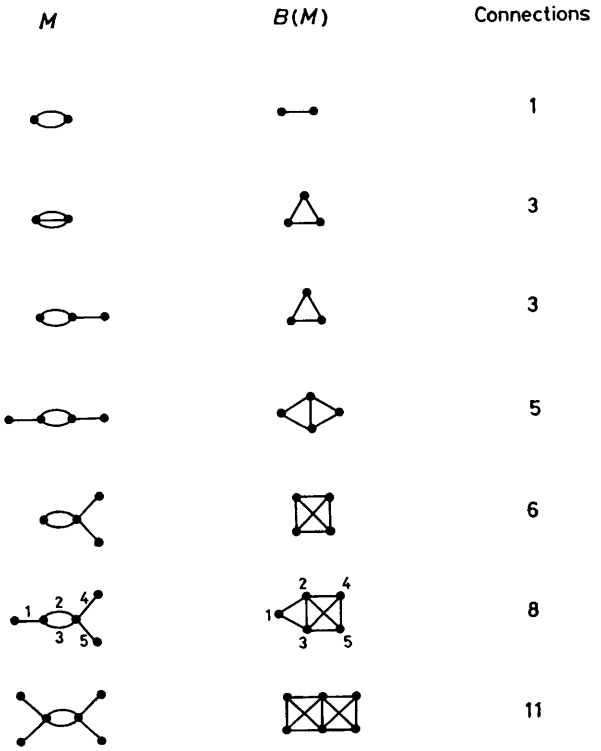

FIGURE 2. Bond graphs of some multiply bonded molecules.

In contrast, while the index of Merrifield and Simmons[9] (the number of open sets in the graph topology) increases with chain length and branching, it decreases with cyclization (see butane, cyclobutane, bicyclobutane, and tetrahedrane). The Randić–Wilkins index[10] (the number of self-avoiding paths) does not change with branching (see the isomeric pentanes). The Bonchev–Trinajstić index[11] (the 'information' on distances) increases with chain length, but decreases with branching and with cyclization, as does the Wiener index[12] (the sum of the distances). The Hosoya index[13] increases with chain length and with cyclization, but decreases with branching. The bond additivity scheme of Randić[14] assigns the same numbers to cyclobutane and tetrahedrane and to cyclohexane and prismane. The Rashevsky index[5] (the 'information' on equivalent points) drops to zero whenever all the atoms are equivalent, no matter how large the molecule.¶ In spite of their limitations, the 'branching' and 'topological' indices mentioned above have been applied to many problems.[1-3,5,9-14]

It may be hoped that the bond graph will aid in these endeavours and new ones.

Added in proof: Dr. W. J. Wiswesser has informed the author that the simple length of his line notation can be used as a rough index of complexity. For the molecules in the Table, his values are (top to bottom): **2**, **2**, **2**; **2**, **5**, **7**; **4**, **4**, **4**; **5**, **11**, **15**; **2**, **2**, **6**.

(*Received*, 24*th February* 1981; *Com.* 209.)

[1] M. Gordon and J. W. Kennedy, *J. Chem. Soc., Faraday Trans.* 2, 1973, **69**, 484; see also J. W. Essam, J. W. Kennedy, M. Gordon, and P. Whittle, *ibid.*, 1977, **73**, 1289.
[2] L. B. Kier and L. H. Hall, 'Molecular Connectivity in Chemistry and Drug Research,' Academic Press, 1976, p. 27—64.
[3] M. Gordon and G. R. Scantlebury, *Trans. Faraday Soc.*, 1964, **60**, 604.
[4] J. B. Hendrickson, *Top. Curr. Chem.*, 1976, **62**, 49 (note that the axis labelled 'molecular complexity' on p. 54 has no numerical scale).
[5] N. Rashevsky, *Bull. Math. Biophys.*, 1955, **17**, 229.
[6] F. Harary, 'Graph Theory,' Addison-Wesley, Reading, Ma., 1969, p. 71.
[7] M. Gordon and W. B. Temple, *J. Chem. Soc. A*, 1970, 729.
[8] W. C. Herndon, M. L. Ellzey, Jr., and K. S. Raghuveer, *J. Am. Chem. Soc.*, 1978, **100**, 2645.
[9] R. E. Merrifield and H. E. Simmons, *Theor. Chim. Acta*, 1980, **55**, 55.
[10] M. Randić and C. L. Wilkins, *Chem. Phys. Lett.*, 1979, **63**, 332.
[11] D. Bonchev and N. Trinajstić, *J. Chem. Phys.*, 1977, **67**, 4517.
[12] H. Wiener, *J. Am. Chem. Soc.*, 1947, **69**, 17.
[13] H. Hosoya, *Bull. Chem. Soc. Jpn.*, 1971, **44**, 2332.
[14] M. Randić, *J. Am. Chem. Soc.*, 1975, **97**, 6609.