CHEMICAL GRAPHS. PART 48.[1]
TOPOLOGICAL INDEX J FOR HETEROATOM-CONTAINING MOLECULES
TAKING INTO ACCOUNT PERIODICITIES OF ELEMENT PROPERTIES

Alexandru T. Balaban

Polytechnic Institute, Organic Chemistry Department,
Splaiul Independentei 313, 76206 Bucharest, Roumania

Abstract. _Two different approaches are presented for the calculation of_
topological index J by taking into account the chemical nature of elements i.
In the first system, relative electronegativities X_i _are calculated semi-_
empirically for 14 atoms with carbon as standard ($X_C = 1$). In the alternative
system, relative covalent radii Y_i _for the same 14 atoms belonging to groups_
3 - 7 for the Periodic System are calculated semiempirically, again with $Y_C = 1$.
Equations for X_i _and_ Y_i _have two parameters : the atomic number_ Z_i, _and the_
number G_i _of the group of the Periodic System (short form), leading to periodic_
variations of X_i _or_ Y_i _vs._ Z_i. _The distance sums are multiplied by_ X_i _or_ Y_i
and then index J is calculated normally.

The topological index J is a number[2-4] associated with a hydrocarbon molecule, and is defined by the formula (1) :

$$J = \frac{q}{p - q + 2} \sum_{\text{edges } ij}' (d_i d_j)^{-1/2} \tag{1}$$

The number of edges and vertices in the hydrogen depleted graph are p and q, respectively. The distance sum d_i (sometimes also called distance degree) for vertex i is a graph invariant and is the sum of topological distances from vertex i to all other vertices in the graph. In formula (1), the summation is over all edges i-j of the hydrogen depleted graph. The presence of double or triple bonds may be taken into account either[2-4] by means of the bond order, or[5] by considering in the summation a double bond as two single bonds, and a triple bond as three single bonds. Computer programs were devised for J.[3,5] The degeneracy of J is the lowest from any single topological index devised so far, as it was established both by us,[4] and by other authors.[6] Interestingly, for many infinite graphs, J attains[7] a finite limit, e.g. for a long, linear alkane J tends towards $\pi = 3.14159$.

A mathematical treatment of the branching problem led Bertz to the conclusion[8] that among all topological indices, J is the only one which orders alkanes in a similar manner to Bertz's graph derivatives.

The presence of heteroatoms in molecules requires the development of special parameters. One approach was developed by Trinajstić and coworkers.[9] These authors defined the diagonal elements d_{ii} in the distance matrix $\underline{\underline{D}}$ as follows :

$$d_{ii} = 1 - 6/Z_i \tag{2}$$

where Z_i is the number of all (valence and inner shell) electrons in atom i. A few values of d_{ii} for various elements are : zero for C ; 0.143 for N ; 0.25 for O ; 0.625 for S ; 0.333 for F ; 0.647 for Cl ; 0.6 for P. Off-diagonal elements in $\underline{\underline{D}}$ are defined as :

$$d_{ij} = \sum_q k_q = \sum_q \frac{36}{b_q Z_i Z_j} \tag{3}$$

where the summation is over all q adjacencies. The bond order b_q is 1 for single bonds, 2 for double bonds and 3 for triple bond, as defined initially by Balaban.[2,3] This leads for the k_q parameter to values such as 0.5 for a C=C bond, 0.67 for an aromatic CC bond, 0.429 for a C=N bond, 0.571 for an aromatic CN bond, 0.75 for C-O bonds, 0.375 for a C=O bond, etc. It should be observed that although heteroatoms may thus be nicely taken into account there is no periodicity which should be an important chemical parameter. Instead, the k_q parameter increases steadily with Z_i.

Another method, due to Basak et al.,[10] makes use of information theory and partitions vertex into classes according to their chemical nature, but without regard to chemical periodicity.

In an approach proposed by Kier and Hall,[11] Randić's molecular connectivity[12] was modified to account for the presence of heteroatoms. Although this method takes into account the number of valence electrons, it does not lead to a rational system incorporating chemical periodicity because for heavier atoms (sulfur, halogens) it uses empirical values based on correlations with molar refractions.

We propose here a different approach in order to take explicitly into account the periodicity of chemical properties for heteroatoms.

Periodicity of heteroatom electronegativities.

According to published data,[13] the electronegativities S of selected Main Group atoms from the first four rows of the Periodic System, recalculated by Sanderson on the Pauling's scale with fluorine having the value 4, sodium 0.56 and hydrogen 2.592, are presented in TABLE 1.

On dividing these values by the carbon electronegativity (2.275) we may obtain non-dimensional relative electronegativities with carbon as standard, which might be tabulated and used empirically as such. We prefer, however, to adopt a semiempirical approach, namely to calculate by a two-parametric linear regression new, linearly adjusted, relative electronegativities. The two parameters for each atom j are Z_j (atomic number) and G_j (the number of the group in Mendeleev's short form of the Periodic System).

TABLE 1. Data[11] for electronegativities S and covalent radii R of 14 elements, on which our calculations are based.

	G	3	4	5	6	7
1st Row	Atom	B	C	N	O	F
	Z	5	6	7	8	9
	S	2.275	2.746	3.194	3.654	4.000
	R(pm)	82.2	77.2	73.4	70.2	68.1
2nd Row	Atom		Si	P	S	Cl
	Z		14	15	16	17
	S	-	2.138	2.515	2.957	3.975
	R(pm)		116.9	110.7	104.9	99.4
3rd Row	Atom			As	Se	Br
	Z			33	34	35
	S	-	-	2.816	3.014	3.219
	R(pm)			119.4	116.7	114.2
4th Row	Atom				Te	I
	Z				52	53
	S	-	-	-	2.618	2.778
	R(pm)				136.0	133.3

We obtain the following electronegativities $S_{j,calc}$ and relative electronegativities X_j, based on the calculated value : S (carbon) = 2.629 :

$$S_{j,calc} = 1.1032 - 0.0204Z_j + 0.4121G_j \qquad (4)$$

$$X_j = 0.4196 - 0.0078Z_j + 0.1567G_j \qquad (5)$$

with correlation coefficient $r^2 = 0.82$. A much better correlation coefficient of $r^2 = 0.997$ would be obtained together with a slightly different equation if only the nine elements of the first two Rows would be considered ; however, this would entail loss of generality, therefore we prefer to use the above equation. Results are presented in TABLE 2.

TABLE 2. Calculated relative electronegativities X and relative covalent radii Y for 14 elements to be used for topological indices J_{het}^{X} and J_{het}^{Y}, respectively

G =		3			4			5						7	
Row	Y	Atom	X	Y	Atom	X	Y	Atom	X	Y	Atom	X	Y	Atom	X
1	1.038	B	0.851	1.000	C	1.000	0.963	N	1.149	0.925	O	1.297	0.887	F	1.446
2		-		1.128	Si	0.937	1.091	P	1.086	1.053	S	1.235	1.015	Cl	1.384
3		-			-		1.379	As	0.946	1.341	Se	1.095	1.303	Br	1.244
4		-			-			-		1.629	Te	0.954	1.591	I	1.103

For calculating the new topological index with heteroatoms (J_{het}) we multiply the distance sum d_j of atom j by X_j, and we apply then the usual formula for J.

Periodicity of heteroatom covalent radii.

Covalent radii R_j of the same 14 atoms are presented (in picometers) in TABLE 1 selected from Sanderson's book.[13]

A similar biparametric correlation affords calculated $R_{j,calc}$ data and relative covalent radii Y_j based on calculated R (carbon) = 87.126 pm, according to equations :

$$R_{j,calc} = 97.4989 + 1.3898Z_j + 4.6779G_j \qquad (6)$$

$$Y_j = 1.1191 + 0.0160Z_j - 0.00537G_j \qquad (7)$$

with a correlation coefficient $r^2 = 0.80$. Again, a much better $r^2 = 0.993$ value would be obtained with slightly different equations by neglecting the last five elements from the 3rd and 4th Rows, but we prefer to include all 14 elements. Results are presented in TABLE 2.

Discussion of the results.

It may be seen from TABLE 2 that within the First-Row atoms from boron to fluorine, the relative calculated electronegativity X increases from 0.851 to 1.446, whereas the relative calculated covalent radius Y decreases from 1.038 to 0.887. A similar but less abrupt change is observed for all other Rows.

On the other hand, within a given group of the Periodic System, the relative calculated covalent radius Y increases with increasing Row number, while the relative calculated electronegativity X decreases.

With respect to carbon for which $X = Y = 1$, elements B, Si, and As have $Y > 1$ and $X < 1$, elements N, O, and F have $Y < 1$ and $X > 1$, while all other chalcogens and halogens have both $X > 1$ and $Y > 1$.

Examples for the topological index calculated on the basis of either relative electronegativities (J^X_{het}) or relative covalent radii (J^Y_{het}) for the four isomeric amines $C_4H_{11}N$ are presented in Fig. 1.

d_i				
$X_N d_N$	10.341	9.192	6.894	5.745
$Y_N d_N$	8.667	7.704	5.778	4.815
J^X_{het}	2.5030	2.4971	2.4540	2.4057
J^Y_{het}	2.5499	2.5516	2.5638	2.5775

Fig.1

It may be seen that distinct values are obtained for all cases. If steric interactions prevail, then J^Y_{het} is smallest for the primary amines (isobutyl-amine < s-butylamine), intermediate for the secondary amine (isopropylmethyl-amine) and maximal for the tertiary amine (ethyldimethylamine) in agreement with the fact that nitrogen has a lower covalent radius than carbon.

The difference between J^Y_{het} and J^X_{het} increases in the same order, because the factors X or Y affect more strongly the J_{het} value when they multiply the smaller sums of the central atoms.

If electronegativity is the more important parameter, then J^X_{het} varies exactly in the reverse order, with the tertiary amine having the lowest, and the primary amines the highest value.

A final example is constituted by the lower alcohols $C_nH_{2n+1}OH$ (TABLE 3).

TABLE 3. Topological indices for the lower alcohols.

n	Alcohol	Steric J_{het}^Y	Electroneg. J_{het}^X
2	EtOH	1.6654	1.5986
3	nPrOH	1.9989	1.9000
3	iPrOH	2.3546	2.2294
4	nBuOH	2.2096	2.1323
4	iBuOH	2.5608	2.4728
4	sBuOH	2.5648	2.4625
4	tBuOH	3.0536	2.9314

It may be seen that both J^Y and J^X increase with increasing branching, leading to the same ordering : nPrOH < iPrOH ; nBuOH < tBuOH. However, isobutanol and sec-butanol are ordered differently : iBuOH < sBuOH via J^Y but iBuOH > sBuOH via J^X. One can also note that the difference between J^Y and J^X increases for each n value in TABLE 3 from top to bottom, i.e. from primary to secondary and to tertiary alcohols, as it was observed above for amines.

In agreement with earlier papers[2-5] one can consider, in addition to the presence of heteroatoms, other factors such as : (i) multiple bonding, either by means of bond orders[2,3] or of counting independently each bond,[5] leading to different results in these two versions ; (ii) the occurence of free valencies for intercomparisons of molecular fragments.[5]

References

1. Part 47, A.T. Balaban and C. Deleanu, *Rev. Roum. Chim.*, (in press).
2. A.T. Balaban, *Chem. Phys. Lett.*, 80, 399 (1982).
3. A.T. Balaban, *Pure Appl. Chem.*, 55, 189 (1983).
4. A.T. Balaban and L.V. Quintas, *Math. Chem.*, 14, 213 (1983).
5. A.T. Balaban and P. Filip, *Math. Chem.*, 16, 163 (1984).
6. M. Razinger, J.R. Chrétien and J.E. Dubois, *J. Chem. Inf. Comput. Sci.*, 25, 23 (1985).
7. A.T. Balaban, N. Ionescu-Pallas and T.-S. Balaban, *Math. Chem.*, 17, 121 (1985).

8. S.H. Bertz, *Discrete Appl. Math.*, (to appear).

9. M. Barysz, G. Jashari, R.S. Lall, V.K. Srivastava and N. Trinajstić, in *"Chemical Applications of Graph Theory and Topology"* (ed. R.B. King), Elsevier, Amsterdam, 1983, p. 222.

10. S.C. Basak, D.P. Gieschen, V.R. Magnuson and D.K. Harris, *ICRS Med. Sci.*, 10, 619 (1982) ; S.C. Basak, A.B. Roy and J.J. Ghosh, in Proc. 2nd Int. Conf. on Mathematical Modeling (eds. X.J.R. Avila, R. Bellman, V.L. Luke and A.K. Riegler), vol. 2, p. 851, Univ. of Missouri, Rolla, 1980 ; A.B. Roy, S.C. Basak, D.K. Harris and V.R. Magnuson, in *"Mathematical Modelling in Science and Technology"*, Proc. 4th Int. Conf., (eds. X.J.R. Avila, R.E. Kalman, A.I. Liapis and E.Y. Rodin), Pergamon Press, New York, 1984, p. 745.

11. L.B.Kier and L.H. Hall, *"Molecular Connectivity in Chemistry and Drug Research"*, Academic Press, New York, 1976.

12. M. Randić, *J. Am. Chem. Soc.*, 97, 6609 (1975).

13. R.T. Sanderson, *"Polar Covalence"*, Academic Press, New York, 1983, p. 41.