On Calculation of the Detour Index

István Lukovits*,‡ and Marko Razinger†,§

Central Research Institute for Chemistry, H-1525 Budapest, POB 17, Hungary, and National Institute of Chemistry, SLO-1000 Ljubljana, Hajdrihova 19, Slovenia

Received June 18, 1996⊗

An algorithm for the detection of the longest path between any two vertices of a graph was proposed. The method was used to calculate the detour index of fused bicyclic structures. Analytical formulas for the detour index of fused bicyclic structures were derived, and it was shown the these can only be third-order polynomials.

INTRODUCTION

The detour index $w(G)$, where G denotes the underlying graph G, has been introduced by Amić and Trinajstić[1] and by John,[2] recently. $w(G)$ is a graph invariant and is an analogue of the well-known Wiener index[3] $W(G)$.

$$w(G) = \sum \Delta_{ij} \tag{1}$$

where Δ_{ij} denotes the length of the longest path between vertices (atoms) i and j and the summation has to be performed over all pairs of vertices $1 \leq i < j \leq N$ where N denotes the total number of vertices. As an example let us consider the hydrogen suppressed graph of cyclobutane (Figure 1a). The values of the entries of the detour matrix are as follows:

$$\Delta_{12} = 3, \quad \Delta_{13} = 2, \quad \Delta_{23} = 3, \quad \Delta_{14} = 3,$$
$$\Delta_{24} = 2, \quad \Delta_{34} = 3$$

The matrix is symmetrical, that is, $\Delta_{ij} = \Delta_{ji}$, and all diagonal entries are equal to zero. The detour index of cyclobutane is therefore $w(\text{cyclobutane}) = 3 + 2 + 3 + 3 + 2 + 3 = 16$.

The detour index is equivalent with the Wiener index $W(G)$ in acyclic structures,[3] where only one path exists between every pair of atoms. In cyclic structures, however, $w(G)$ and $W(G)$ are different. Another variant of $W(G)$, being equivalent to $W(G)$ in acyclic structures and nonequivalent to it in cyclic structures, is the Szeged index.[4] Randić has extended the algorithm devised by Wiener for adjacent vertices to all pairs of vertices and obtained the "hyper-Wiener" index.[5] A quite new family of indices also related to $W(G)$ has been defined by Diudea.[6] Mathematical properties and applications of the Wiener index were summarized in reviews published recently.[7−9] It has to be noted that the Rouvray index,[10] proposed independently of W, is equal to $2W(G)$.

It was found that $w(G)$ is useful in quantitative structure−property (QSPR) studies, especially if it is combined with the Wiener index.[11] However, the application of detour index is problematic, because no algorithm has been published to compute $w(G)$ so far. Quite recently, Trinajstić has reported the successful construction of such an algorithm.[12] In fact the construction of an efficient algorithm would be equivalent to the solution of the famous NP-complete problem.

In this paper we propose an algorithm that can be applied for the computation of $w(G)$. The algorithm has been used to generate the detour indices for various condensed bicyclic structures, which in turn were used to derive formulas for $w(G)$ of such structures. Formulas for graph invariants may be used to solve the "graph reconstruction" problem[13,14] numerically. For this the numerical values of invariants for a given class of structures have to be generated and checked with structures corresponding to the prescribed value. In this paper it was shown that $w(G)$ is polynomial of the sizes of strings (see definition below) the order of which is not higher than three, provided that such formulas can be derived. For fused bicyclic structures, which are not cutpoint graphs, it has been proven[15] that such polynomials can be derived.

THE ALGORITHM

Only hydrogen suppressed graphs of hydrocarbons will be considered, the degree of their vertices being less or equal to four. Expressions "atom" and "vertex", "chemical bond" and "edge", and "valence" and "degree" will be used interchangeably, hereafter. Path starting at vertex i and ending at vertex j, where $i < j$, will be denoted by $P(i \rightarrow j)$.

The construction of the sequence of vertices will be illustrated for atoms 2 and 7 of 2-ethylnaphthalene. The algorithm is a "back-track" procedure.[16,17] In our example the procedure starts at atom 2 (Figure 1b) which is the first member of a sequence. The first neighbors are vertices 1 and 3. A vertex with a lower index (in our case vertex 1) will be considered first. Branching vertices with neighbors to be examined will be denoted by an asterisk (*). Branching vertices with no more neighbors to be examined will be denoted by two asterisks. The procedure "back-tracks" to the nearest branching vertex with noninspected neighbors if 1. an endpoint is encountered; 2. an atom already appearing in the sequence is encountered; and 3. if the target atom j (in our example vertex 7) is encountered. Once a path $P(2 \rightarrow 7)$ has been found, it will be typed in **boldface**. Sequences which terminate without containing the target index will be underlined. The procedure stops if there is no branching atom with noninspected neighbors left. Note that the starting vertex, unless it is an endpoint, will be treated

† Deceased on July 31, 1996.
‡ Central Research Institute for Chemistry.
§ National Institute of Chemistry.
⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1997.

S0095-2338(96)00034-0 CCC: $14.00

as a branching vertex. The consecutive lines illustrate the evolution of the sequences.

2*

2*, 1*

2*, 1*, 10

2*, 1*, 10, 9*

2*, 1*, 10, 9*, 4*

2*, 1*, 10, 9*, 4*, 3

2*, 1*, 10, 9*, 4*, 3, 2*

2*, 1*, 10, 9*, 4**, 5

2*, 1*, 10, 9*, 4**, 5, 6

2*, 1*, 10, 9*, 4, 5, 6, 7**

2*, 1*, 10, 9**, 8

2*, 1*, 10, 9, 8, 7**

2*, 1**, 11

2*, 1**, 11, 12

2**,3

2**, 3, 4*

2**, 3, 4*, 5

2**, 3, 4*, 5, 6

2, 3, 4*, 5, 6, 7**

2**, 3, 4**, 9*

2**, 3, 4**, 9*, 8

2, 3, 4**, 9*, 8, 7**

2**, 3, 4**, 9**, 10

2**, 3, 4**, 9**, 10, 1*

2**, 3, 4**, 9**, 10, 1*, 2

2**, 3, 4**, 9**, 10, 1**, 11

2**, 3, 4**, 9**, 10, 1**, 11, 12

The longest path between vertices 2 and 7 is therefore 2, 1, 10, 9, 4, 5, 6, 7 and $\Delta_{27} = 7$.

The longest path may be stored in matrix **A**. Matrix **A** has $N(N-1)/2$ rows. The ath ($a = 1, ..., N(N-1)/2$) row of matrix **A** contains path P($i \rightarrow j$). The serial number a is composite index and is uniquely determined by i and j:

$$a = (j^2 - 3j + 2)/2 + i \qquad (2)$$

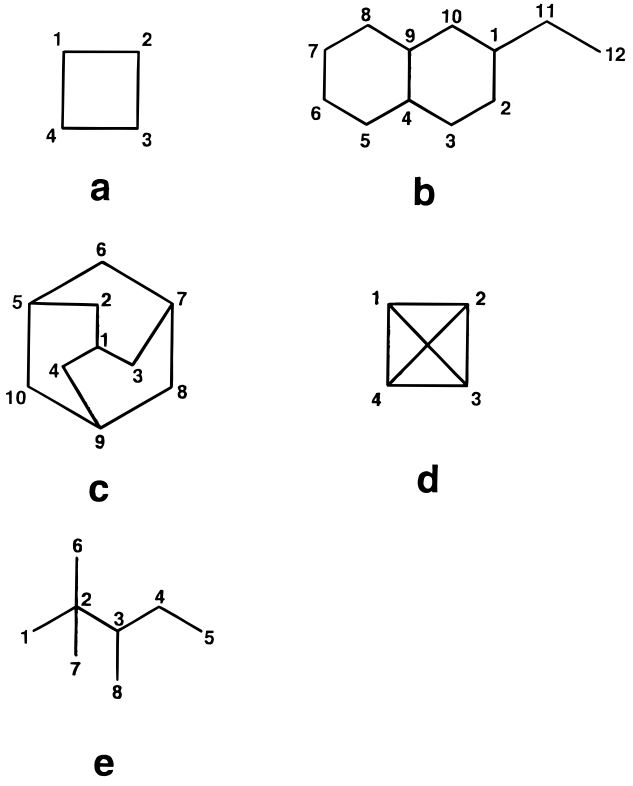

Figure 1.

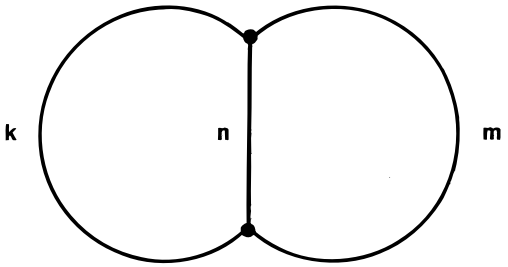

Figure 2.

In this way the position of each path in matrix **A** unambiguously defines the first and the last atom in the path, and, therefore, it is not necessary to store i and j. The longest possible path contains $N - 2$ vertices, and the dimension of matrix **A** is $N(N-1)/2 \times (N-2)$. The composite index a can be decomposed into i and j by using the following algorithm:

1. set $j = 0$
2. $j = j + 1$
3. if $j(j-1)/2 < a$, go to step 2
4. $i = a - (j-1)(j-2)/2$

For $i = 2$ and $j = 7$, $a = 17$.

The calculated values of the detour index in structures a, b, c, d, and e (Figure 1) are 16, 514, 324, 18, and 63, respectively. We tested the program for 37 structures (Figure 2 in ref 1). In most cases the calculated values of $w(G)$ were identical with the reported figures, but for structures 15, 26, and 37 the correct values are 29, 91, and 69, respectively.

A THEOREM

A string is a subgraph starting with an endpoint or a branching atom and ending with an endpoint or a branching atom.[14] The degrees of vertices between these (if any) are equal to two. Note that letters k, m, and n (Figure 2) denote both the strings themselves and the number of vertices in

Table 1. Table of Strings Used to Derive Formulas for the Detour Index

case 1				case 2				case 3				case 4			
k	*m*	*n*	detour index	*k*	*m*	*n*	detour index	*k*	*m*	*n*	detour index	*k*	*m*	*n*	detour index
2	2	2	1	5	5	4	315	8	5	3	584	9	6	3	951
4	2	2	16	7	5	2	344	10	7	5	1996	11	8	5	2778
6	4	2	168	9	7	4	1412	8	7	7	1926	9	8	7	2687
4	4	2	67	9	9	8	3563	6	5	5	552	11	10	5	3746
6	2	2	63	7	7	6	1345	4	3	3	62	5	4	3	161
8	4	2	337	9	5	4	919	8	7	5	1380	13	10	5	4904
8	2	2	160	7	3	2	165	8	3	3	319	13	6	5	2737
10	4	4	903	11	9	4	2836	10	9	3	2092	11	8	3	2085
10	6	2	975	11	5	4	1386	10	5	5	1365	7	6	5	904
8	6	2	608	9	7	6	1962	8	7	3	956	9	4	3	576
6	4	4	315	5	5	2	171	6	5	3	330	15	8	5	4853
6	6	4	566	7	7	4	927	10	9	5	2782	15	10	3	5054
8	6	4	925	9	9	4	2043	12	7	3	2069	15	6	3	3658
8	8	2	987	11	11	2	2969	12	3	3	908	13	10	3	3881
10	8	6	2730	13	5	4	1987	12	5	5	1959	17	6	5	4764
8	4	4	560	13	7	4	2802	14	5	3	2024	17	8	3	4989
6	6	6	875	11	7	6	2725	10	7	7	2690	11	8	7	3629
8	8	4	1410	13	7	2	2116	12	7	5	2761	13	12	3	5085
8	6	6	1350	3	3	2	17	10	9	7	3626	11	10	7	4763
8	8	8	2632	9	9	6	2727	10	9	9	4640	11	10	9	5966

the strings. The total number of atoms of structure depicted in Figure 2 is therefore $N = k + m + n - 4$.

It will be shown that $w(K_N)$, where K_N denotes a complete graph with N vertices is a third-order polynomial in terms of N. This is because for K_N, $w(K_N) = N(N-1)^2/2$. Since there is no other graph for which the length of every longest path between each pair of vertices is equal to $N - 1$, $w(K_N)$ has the highest numerical value of all graphs with N vertices. But N is the sum of the sizes of strings minus a "correction" factor. Therefore we can state our theorem:

> If for $w(G)$ a polynomial in terms of sizes of the strings could be found, then the order of this polynomial would not be higher than three.

It was shown that for fused bicyclic structures such polynomial can be constructed.[15]

APPLICATION: FORMULAS FOR FUSED BICYCLIC STRUCTURES

The present algorithm has been used to compute $w(G)$ for several fused bicyclic structures (Figure 2). For condensed bicyclic structures $w(G)$ may be expressed in terms of sizes of the strings as follows:

$$w(G) = Ak^3 + Bm^3 + Cn^3 + Dk^2m + Ek^2n + Fm^2n + Gkm^2 + Hkn^2 + Imn^2 + Jkmn + Kk^2 + Lm^2 + Mn^2 + Nkm + Okn + Pmn + Rk + Sm + Tn + U \quad (3)$$

Coefficients A, B, ..., U are unknown, but they can be obtained if for 20 various structures $w(G)$ were known and a set of simultaneous linear equations is solved for coefficients A, B, ..., U. This technique has already been used for derivation of formulas for the Wiener index.[14]

The formulas for simple cycles with N even or odd are not equivalent.[11] For fused bicycles eight ($2^3 = 8$) cases had to be considered, but there are four equivalencies within these classes and we obtain the following four cases:

1. k, m, and n are even *or* k, m, and n are odd,
2. k and m are odd and n is even *or* k and m are even and n is odd,
3. k is even, m and n are odd *or* k is odd and m and n are even,
4. k and n are odd and m is even *or* k and n are even and m is odd.

The coefficients cannot depend on the actual values of k, m, and n. In practice, however, many combinations of k, m, and n often yield a matrix of polynomials (k^3, m^3, ..., k^2m, ..., k, m, n, constant) which is singular. Therefore we listed those k, m, and n (Table 1) for which the polynomials do not form a singular matrix. Note that the first subcases were considered for each case only, but the formula remains valid for the second subcase as well. The equations were solved by using a standard matrix inversion routine. The final equations of w(G) are as follows:

Case 1:

$$w(G) = (9k^3 + 7m^3 + 8n^3 + 27k^2m + 18k^2n + 18m^2n + 33km^2 + 30kn^2 + 18mn^2 + 60kmn - 102k^2 - 102m^2 - 120n^2 - 264km - 216kn - 204mn + 468k + 500m + 484n - 672)/24 \quad (4)$$

Case 2:

$$w(G) = (9k^3 + 7m^3 + 8n^3 + 27k^2m + 18k^2n + 18m^2n + 33km^2 + 30kn^2 + 18mn^2 + 60kmn - 102k^2 - 102m^2 - 120n^2 - 264km - 216kn - 204mn + 474k + 506m + 484n - 708)/24 \quad (5)$$

Case 3:

$$w(G) = (9k^3 + 7m^3 + 8n^3 + 27k^2m + 18k^2n + 18m^2n + 33km^2 + 30kn^2 + 18mn^2 + 60kmn - 102k^2 - 102m^2 - 120n^2 - 264km - 216kn - 204mn + 471k + 503m + 490n - 690)/24 \quad (6)$$

Case 4:

$$w(G) = (9k^3 + 7m^3 + 8n^3 + 27k^2m + 18k^2n + 18m^2n + 33km^2 + 30kn^2 + 18mn^2 + 60kmn - 102k^2 - 102m^2 - 120n^2 - 264km - 216kn - 204nm + 477k + 497m + 490n - 690)/24 \quad (7)$$

CONCLUSIONS

The proposed algorithm is not effective generally, that is it does not solve the NP-complete problem. The efficiency of the algorithm is similar to that proposed by Xu.[18] For most chemical applications, however, the speed of the procedure will be sufficient.

The formulas for the four subcases differ only in the first- and zero-order terms, similarly to the formulas derived for the Wiener index.[14] In analogy with the formulas obtained for $W(G)$ coefficients (eqs 4−7) of the third- and first-order terms are positive, whereas the coefficients of the second- and zero-order terms are negative.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Science and Technology of Slovenia for financial support (Grant. No. J1-6161). I.L. thanks the National Institute of Chemistry, Ljubljana, for the hospitality during his visit.

REFERENCES AND NOTES

(1) Amić, D.; Trinajstić, N. On the Detour Matrix. *Croat. Chem. Acta* **1995**, *68*, 53−62.
(2) John, P. On the computation of the Wiener index of selected Delta-dimensional lattice structures. *MATCH Commun. Math. Chem.* **1995**, *32*, 207−219.
(3) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17−20.
(4) Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I. Dömötör, Gy. The Szeged Index and an Analogy with the Wiener Index. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547−550.
(5) Randić, M. Novel Molecular Descriptor for Structure−Property Studies. *Chem. Phys. Lett.* **1993**, *211*, 478−483.
(6) Diudea, M. V. Walk Numbers $^{e}W_{M}$: Wiener-Type Numbers of Higher Rank. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 535−540.
(7) Gutman, I.; Yeh, Y. N.; Lee, S. L.; Luo, Y. L. Some recent results in the theory of the Wiener number. *Indian J. Chem.* **1993**, *32A*, 651−661.
(8) Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener Index: Development and Applications. *Croat. Chem. Acta* **1995**, *68*, 105−129.
(9) Zhu, H. Y.; Klein, D. J.; Lukovits, I. Extensions of the Wiener-Number. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 420−428.
(10) Rouvray, D. The Value of Topological Indices in Chemistry. *MATCH* **1975**, *1*, 125−134.
(11) Lukovits, I. The Detour Index. *Croat. Chem. Acta* **1996**, *69*, 873−882.
(12) Nikolić, S.; Trinajstić, N.; Jurić, A.; Mihalić, Z. On the Detour Matrix and the Detour Index for Weighted Graphs. *Croat. Chem. Acta* Submitted for publication.
(13) Kier, L. B.; Hall, L. H.; Frazer, J. W. Design of Molecules from Quantitative Structure - Activity Relationship Models. I. Information Transfer between Path and Vertex Degree Counts. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 143−147.
(14) Lukovits, I. General Formulas for the Wiener Index. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 503−507.
(15) Linert, W.; Lukovits I. Formulas for the Hyper-Wiener and Hyper-Detour Indices of Fused Bicyclic Structures. *MATCH* Submitted for publication.
(16) Ray, L. C.; Kirsch, R. A. Finding Chemical Records by Digital Computers. *Science* **1957**, *126*, 814−819.
(17) Barnard, J. M. Substructure Search Methods: Old and New. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 532−538.
(18) Xu, J. GMA: A Generic Match Algorithm for Structural Homomorphism, Isomorphism, and Maximal Common Substructure Match and Its Applications. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 25−34.

CI960034J