

D_{MAX} – Matrix of Dominant Distances in a Graph^{*}

Milan Randić^{**}

National Institute of Chemistry, Ljubljana Hajdrihova 19, Slovenia
mrandic@msn.com

(Received May 23, 2012)

Abstract

We are introducing a novel distance-type matrix for graphs, referred to as D_{MAX} , which is constructed from the distance matrix of a graph by selecting for each row and each column only the largest distances. This matrix can be viewed as opposite to the adjacency matrix, which can be constructed from the distance matrix by selecting for each row and each column only the smallest distances, which correspond to adjacent vertices. We have illustrated the novel matrix for linear n -alkanes and n -star graphs, as well as the nine isomers of heptane. In the case of isomers, the matrix appears very sensitive on the branching pattern of molecular graphs. Therefore it may be of interest for characterizing of molecular shapes. On the other hand, in view that selected invariants of D_{MAX} for structurally related molecules could be significantly different, it appears that such invariants of D_{MAX} can be of interest in testing graphs for isomorphism.

1. Introduction

Graphs have been of interest in chemistry for quite a while, primarily because selected graph invariants can serve as molecular descriptors. Graph invariants are any of mathematical properties of graphs that are independent of the numbering of graph vertices or of the geometrical representation of a graph. One class of graph invariants follows as matrix invariants

^{*} This contribution is dedicated to Professor Ivan Gutman, on the occasion of his 65th anniversary, for his scientific accomplishments as outstanding pioneer of Mathematical Chemistry, for his contributions to Graph Theory and Chemical Graph Theory, and his leadership as the current Editor of *MATCH Communications in Mathematical and Computer Chemistry*.

^{**} Emeritus, Dept. of Mathematics & Computer Science, Drake University, Des Moines, IA, USA

from representations of graphs by matrices. The well-known and oldest graph matrices are: the adjacency matrix, introduced by Poincaré for characterizing of labyrinths [1]; and the distance matrix, introduced by Frank Harary [2]. In more recent time, with the growth of Chemical Graph Theory, additional graph matrices have been introduced, like: the Wiener matrix [3-7]; the path-Wiener matrix [8]; the Szeged matrix [9, 10] and the revised Szeged matrix [11]; the Distance/Distance matrix [12]; resistance-distance matrix [13, 14]; novel graph distance matrix based on viewing rows as points in n -dimensional space [15], as well as matrices the elements of which are not numbers, but subgraphs, such as the double invariants matrix [16] and the Ulam subgraph matrix [17]. For a more complete review on matrices in chemistry we direct readers to the book of Janežič *et al.* [18].

In this contribution we would like to introduce a novel distance-type graph matrix, which is characterized by its simple conceptual content as well as the simple constructional procedure. What makes this matrix in some respect different from most of the above mentioned matrices is that it is sensitive to the pattern of branching in a graph, at least for the tree graphs (acyclic graphs) that we consider in this contribution. So two apparently similar graphs can nevertheless have visibly different matrices, which will consequently have a number of fairly different graph invariants. While such matrices and their invariants are therefore to of limited and little interest in structure-property and structure-activity studies, because similar molecules will have dissimilar molecular descriptors, they may be of considerable interest for characterization of graph shapes and for investigating graph isomorphism.

2. Construction of the D_{MAX} Matrix of a Graph

We will illustrate the construction of the D_{MAX} matrix on the molecular graph of 4-ethyl-2-methyl-heptane. We start by construction of the distance matrix for this graph, which is illustrated in Table 1. In Table 2 we show for each row of the distance matrix only its largest entries (distances). When this matrix is transposed one obtains a portion of the distance matrix showing the largest entries for each column. When one combines the matrix elements which are the largest in each row and in each column of the distance matrix and sets all missing matrix elements in the matrix to be zero, one obtains the D_{MAX} matrix of 4-ethyl-2-methyl-heptane shown in Table 3. Formally the definition of the D_{MAX} matrix is:

Table 3 The D_{MAX} for 2-methyl-4-ethyl-heptane

	1	2	3	4	5	6	7	8	9	10
1	0	0	0	3	4	5	6	0	4	5
2	0	0	0	0	0	0	5	0	0	0
3	0	0	0	0	0	0	4	0	0	0
4	3	0	0	0	0	0	0	3	0	0
5	4	0	0	0	0	0	0	4	0	0
6	5	0	0	0	0	0	0	5	0	0
7	6	5	4	3	0	0	0	6	4	5
8	0	0	0	3	4	5	6	0	4	5
9	4	0	0	0	0	0	4	4	0	0
10	5	0	0	0	0	0	5	5	0	0

of view one can recognize the novel matrix as one of the more elementary graph matrices, ranking in importance probably next to the adjacency matrix and the distance matrix. It seems therefore of considerable interest to investigate the mathematical properties of this novel graph matrix.

3. Reconstruction of a Graph from D_{MAX}

The first and important question is whether the D_{MAX} allows the reconstruction of the graph, because if it does, then this means that there is no loss of information on the graph and that no two graphs can have the same D_{MAX} . Even though D_{MAX} is defined for general graphs (acyclic or cyclic) we will confine our attention mostly to trees. It is known that in case of trees it is sufficient to know all the distances between its terminal vertices. Once the terminal vertices have been identified, one can find all the distances between terminal vertices. According to a theorem of Zaretskii [19] this information suffices for graph reconstruction. The question is whether from D_{MAX} one can extract information on distances between all terminal vertices. Clearly the largest entries in D_{MAX} will identify several terminal vertices, but one needs distances between all terminal vertices.

In the case of 4-ethyl-2-methyl-heptane we can see from Table 3 that matrix elements (1,7) and (7,8) are the largest. This immediately suggests that the vertices 1, 7, 8 are terminal. Moreover, it is not difficult to deduce that graph of 2-methylheptane is a subgraph of the graph of 4-ethyl-2-methyl-heptane graph, which accounts for eight of ten carbon atoms of 2-methyl-4-ethylheptane. This information suffices to construct the part of the adjacency matrix belonging

to vertices 1-8. From this part of attempted reconstruction we found out that the distance between terminal vertices 1 and 2 is not included in D_{MAX} , though it was possible to determine it. Hence, D_{MAX} does not necessarily involve information on all distances between terminal vertices. From the remaining columns and rows of the D_{MAX} matrix of 4-ethyl-2-methyl-heptane one finds that matrix elements (1,10), (7,10) and (8,10) are the largest entries. This points that vertex 10 is at the same distance vertices 1, 7, and 8. The only vertex at the same distance from these three vertices is vertex 4, which must be the site of ethyl group. This suffices to identify vertex 4 where the ethyl branch of 4-ethyl-2-methyl-heptane is attached to 1,2-dimethylheptane subgraph to make 4-ethyl-2-methyl-heptane – which completes the reconstruction of this particular graph. This, of course, is not a proof for reconstruction of graph of trees from their D_{MAX} matrix, but illustrates the procedure of reconstruction.

It may be very difficult to prove that D_{MAX} allows reconstruction for a general graph (or tree), but if one can find two non isomorphic graphs having the same D_{MAX} this would be counterexample for reconstruction conjecture, and this would definitely prove that there is loss of information accompanying construction of D_{MAX} . Well, as is discussed in the following article by Randić et al. [20] there is a pair of graphs (trees) on 18 vertices, which have the same D_{MAX} , which ends all speculation of reconstruction of graph from D_{MAX} for a general case.

Let us also point out that D_{MAX} should not be confused with the detour matrix of graphs, introduced by Harary [2], the elements of which record the length of the *longest* distance between each pair of vertices. For comparison, the non-zero elements of the D_{MAX} form a submatrix of the distance matrix, which records the length of the *shortest* path between vertices. The elements of the distance matrix entering into D_{MAX} are the longest paths among the *shortest* path between vertices. In Table 4 we have illustrated D_{MAX} for the three graphs of Fig. 1 which all have the same detour matrix.

Fig. 1 Graphs having the same detour matrix

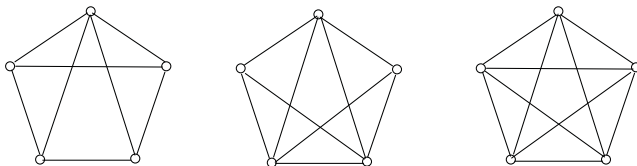


Table 4 The D_{MAX} matrices for three graphs of Fig. 1 and their detour matrix (at right)

0 1 1 1 1	0 1 1 1 1	0 1 1 1 1	0 4 4 4 4
1 0 0 2 0	1 0 1 1 2	1 0 1 1 1	4 0 4 4 4
1 0 0 0 2	1 1 0 1 1	1 1 0 1 1	4 4 0 4 4
1 2 0 0 0	1 1 1 0 1	1 1 1 0 1	4 4 4 0 4
1 0 2 0 0	1 2 1 1 0	1 1 1 1 0	4 4 4 4 0

As we see from Table 4, the three graphs have different D_{MAX} but the same detour matrix shown at the most right. It happened in this illustration that that for the central graph and $K_{5,5}$ (the last graph) D_{MAX} is identical to the distance matrix

4. D_{MAX} Matrices for selected graphs

In this section we will briefly discuss D_{MAX} matrices for several simple graphs, for which one can construct D_{MAX} because of having simple analytical form.

D_{MAX} of n -alkanes

In Table 5 we show D_{MAX} for the molecular graphs of non-branched alkanes propane, butane, pentane, and hexane. It is immediately clear that these matrices have non-zero entries only in the first and the last row and column. Their structure is very simple and it is not difficult to construct D_{MAX} for such n -alkanes, regardless the value of n . If for these matrices one is to add non-zero elements above the main diagonal one obtains for odd size matrices 4, 14, 30, 52,

80, and so on, and for even size matrices 7, 19, 37, 61, 91, and so on. In both cases the second difference between the adjacent members in these sequences is equal to 6, the first differences being 10, 16, 22, 28, . . . and 12, 18, 24, 30, . . ., respectively.

Table 5 The D_{MAX} matrices for graphs of smaller n -alkanes

Propane	Butane	Pentane	Hexane
0 1 2	0 0 2 3	0 0 2 3 4	0 0 0 3 4 5
1 0 1	0 0 0 2	0 0 0 0 3	0 0 0 0 0 4
2 1 0	2 0 0 0	2 0 0 0 3	0 0 0 0 0 3
	3 2 0 0	3 0 0 0 0	3 0 0 0 0 0
		4 3 2 0 0	4 0 0 0 0 0
			5 4 3 0 0 0

D_{MAX} matrices of n -star graphs

In Table 6 we show the D_{MAX} matrix for graphs of small star graphs. As one can see, in this case D_{MAX} matrices are the same as the corresponding distance matrices of these graphs, because all vertices are at the same distance one from each other, except for the central vertex which is adjacent to all other vertices.

Table 6 The D_{MAX} matrices for smaller star graphs

3-star	4-star	5-star
0 1 1 1	0 1 1 1 1	0 1 1 1 1 1
1 0 2 2	1 0 2 2 2	1 0 2 2 2 2
1 2 0 2	1 2 0 2 2	1 2 0 2 2 2
1 2 2 0	1 2 2 0 2	1 2 2 0 2 2
	1 2 2 2 0	1 2 2 2 0 2
		1 2 2 2 2 0

D_{MAX} of the complete graphs K_n

Because each vertex in a complete graph K_n is adjacent to all other $n-1$ vertices, the distance matrix (D) of the complete graphs is identical to their adjacency matrix (A), which has all entries 1 except on the main diagonal. Thus all entries in each row and each column are the maximal entries and the D_{MAX} is the same as the adjacency matrix or the distance matrix of such graphs.

D_{MAX} of the complete bipartite graphs $K_{m,n}$

Vertices of complete bipartite graphs $K_{m,n}$ form two classes. Vertices of the two classes are at distance 1, while distances of vertices that belong to different class are separated by distance of 2, as can be seen from Table 7 (left) in which, for illustration, we show the distance matrix for graph $K_{3,5}$. From this it follows that the D_{MAX} of a bipartite graph consists of two diagonal blocks of size m and n , respectively, all the elements of which are 2, as is shown in Table 7 (at right).

Table 7 Distance matrix and D_{MAX} of the complete bipartite graphs $K_{3,5}$

Distance matrix of $K_{3,5}$									D_{MAX} matrix of $K_{3,5}$								
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
1	0	2	2	1	1	1	1	1	1	0	2	2	0	0	0	0	0
2	2	0	2	1	1	1	1	1	2	2	0	2	0	0	0	0	0
3	2	2	0	1	1	1	1	1	3	2	2	0	0	0	0	0	0
4	1	1	1	0	2	2	2	2	4	0	0	0	0	2	2	2	2
5	1	1	1	2	0	2	2	2	5	0	0	0	2	0	2	2	2
6	1	1	1	2	2	0	2	2	6	0	0	0	2	2	0	2	2
7	1	1	1	2	2	2	0	2	7	0	0	0	2	2	2	0	2
8	1	1	1	2	2	2	2	0	8	0	0	0	2	2	2	2	0

5. D_{MAX} Matrices for heptane isomers

In this section we will discuss the D_{MAX} matrices for the nine isomers of heptanes, shown in Fig. 2; the D_{MAX} matrices are shown in Table 8. Let us look at the D_{MAX} matrices of heptanes isomers more closely. They certainly appear different if one only counts the number of non-zero matrix elements that they have, which vary between 12 (for 3-methylhexane) and twice that number, 24 (for 3-ethylpentane). In Table 9 in the first numerical column we show the number of non-zero entries above the main diagonal, which are half of the previously mentioned numbers. The adjacent column gives the sum of non-zero entries above the main diagonal. In the case of distance matrix the sum of all entries above the main diagonal (which are necessarily non-zero) give the Wiener number, W , which we have listed for comparison in the last column of Table 9. Because of analogy we will use small letter w as the symbol for the sum of all entries above the main diagonal. Observe that among heptanes isomers there is a pair of isomers (2,2-dimethylpentane, 2,3-dimethylpentane) having the same value of w . In the case of Wiener number the same pair also shows identical $W = 46$. In addition, there is another pair of heptanes isomers (3-ethylpentane, 2,4-dimethylpentane) having the same Wiener number ($W = 48$). It appears therefore that w may have a higher discrimination property than W , but both invariants are not very good for discrimination among graphs, if used alone. Actually, it is well known that Wiener number in this respect is notoriously weak, as can be seen from the fact that among 35 nonane isomers there are 29 isomers having no unique W , and among 72 decane isomers there are 54 isomers having no unique W , which one can verify from a paper of Balaban *et al* [21], where W values are listed for all alkanes having twelve and less carbon atoms.

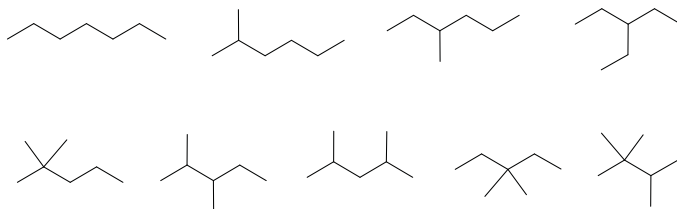


Fig. 2

The nine isomers of *n*-heptane

In the next column of Table 9, column (π), we show the partition of number w into contributions from distances of different length, starting with the maximal distance. Thus the code 1, 2, 2, 2 for n -heptane stand for: $6^1, 5^2, 4^2, 3^2$, that is, among the maximal distance in D_{MAX} matrix of n -heptane we have above the main diagonal one distance of length six and two distances of length five, four, and three. As one can see, all partitions of w for heptanes isomers are different, thus partitioning of w differentiates among heptanes isomers having the same w . The same is also the case with the ordered sequences of row sums, shown in the next column. If all entries in these sequences are added, one obtains $2w$. This offers a useful independent way to check calculations of w .

Table 8 The D_{MAX} matrices for the nine heptanes isomers (the standard numbering of carbon atoms is assumed)

n-heptane							2-methylhexane							3-methylhexane									
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7			
1	0	0	0	3	4	5	6	1	0	0	0	3	4	5	0	1	0	0	0	3	4	5	0
2	0	0	0	0	0	0	5	2	0	0	0	0	0	4	0	2	0	0	0	0	0	4	0
3	0	0	0	0	0	0	4	3	0	0	0	0	0	3	0	3	0	0	0	0	0	3	0
4	3	0	0	0	0	0	3	4	3	0	0	0	0	3	4	3	0	0	0	0	0	0	
5	4	0	0	0	0	0	0	5	4	0	0	0	0	4	5	4	0	0	0	0	0	0	
6	5	0	0	0	0	0	0	6	5	4	3	0	0	5	6	5	4	3	0	0	0	4	
7	6	5	4	3	0	0	0	7	0	0	0	3	4	5	0	7	0	0	0	0	0	4	0

3-ethylpentane							2,2-dimethylpentane							2,3-dimethylpentane									
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7			
1	0	0	2	3	4	3	4	1	0	0	2	3	4	0	0	1	0	0	2	3	4	0	3
2	0	0	0	0	3	0	3	2	0	0	0	0	3	0	0	2	0	0	0	0	3	0	0
3	2	0	0	0	2	0	2	3	2	0	0	0	2	2	2	3	2	0	0	0	2	2	0
4	3	0	0	0	0	0	3	4	3	0	0	0	3	3	4	3	0	0	0	0	3	0	
5	4	3	2	0	0	3	4	5	4	3	2	0	0	4	4	5	4	3	2	0	0	4	3
6	3	0	0	0	3	0	0	6	0	0	2	3	4	0	0	6	0	0	2	3	4	0	3
7	4	3	2	3	4	0	0	7	0	0	2	3	4	0	0	7	3	0	0	0	3	3	0

2,4-dimethylpentane	3,3-dimethylpentane	2,3,3-trimethylbutane
1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
1 0 0 2 3 4 0 0	1 0 0 2 3 4 3 3	1 0 0 2 3 0 0 3
2 0 0 0 0 3 0 3	2 0 0 0 0 0 3 0	2 0 0 0 2 0 0 2
3 2 0 0 0 0 2 2	3 2 0 0 0 0 2 0	3 2 0 0 0 2 2 0
4 3 0 0 0 0 3 0	4 3 0 0 0 0 0 0	4 3 2 0 0 3 3 0
5 4 3 2 0 0 4 0	5 4 3 2 0 0 3 3	5 0 0 2 3 0 0 3
6 0 0 2 3 4 0 4	6 3 0 0 0 3 0 0	6 0 0 2 3 0 0 3
7 4 3 2 0 0 4 0	7 3 0 0 0 3 0 0	7 3 2 0 0 3 3 0

Table 9 The number of non-zero entries above the main diagonal in D_{MAX} of heptane isomers (column #); the sum of non-zero entries above the main diagonal (column w); the quotient of the two, giving the average matrix element (column Q); partition of w coming from distances of decreasing length (column π); ordered row sums (column Σ); and the Wiener number (column W).

	#	w	Q	π	Σ	W
n-heptane	7	30	4.2857	1, 2, 2, 2	18, 18, 6, 5, 5, 4, 4	56
2-methylhexane	8	31	3.8750	0, 2, 3, 3	17, 12, 12, 8, 6, 4, 3	52
3-methylhexane	6	23	3.8333	0, 0, 1, 3, 2	16, 12, 4, 4, 4, 3, 3	50
3-ethylpentane	12	36	3.0000	0, 0, 3, 6, 3	16, 16, 16, 6, 6, 6, 6	48
2,2-dimethylpentane	11	32	2.9091	0, 0, 3, 4, 4	17, 9, 9, 9, 9, 8, 3	46
2,3-dimethylpentane	11	32	2.9091	0, 0, 2, 6, 3	16, 12, 12, 9, 6, 6, 3	46
2,4-dimethylpentane	10	33	3.3000	0, 0, 3, 4, 3	13, 13, 13, 9, 6, 6, 6	48
3,3-dimethylpentane	9	26	2.8889	0, 0, 1, 6, 2	15, 15, 6, 6, 4, 3, 3	44
2,2,3-trimethylbutane	11	28	2.5455	0, 0, 0, 6, 5	11, 11, 8, 8, 8, 6, 4	42

The quotient values in the central column of Table 9, which give the average value of matrix elements of D_{MAX} , represent a novel invariant, which in the case of the distance matrix D , gives that same information as the Wiener number W , in view that the distance matrix has no zero elements outside the main diagonal. It is therefore of considerable interest to see if Q values show some structural regularities.

A way to check this is to see if Q values show a regular dependence on the number of paths of length two and paths of length three, p_2 and p_3 , respectively), which appear to be the case with most topological indices. Randić and Wilkins [22, 23] have shown that the variation of different physicochemical properties among isomers of alkanes, including properties that do not correlate among themselves, such as the heats of formation and the molar refraction, when plotted over a coordinate grid with p_2 and p_3 as the x, y coordinates, show regular variations with the counted p_2 and p_3 of isomers. In Fig. 3 (left) we show the diagram in which the nine isomers of heptanes have been placed over a (p_2, p_3) coordinate system and in Fig. 3 (right) the molecular diagrams of isomers have been replaced by the corresponding Q value of each isomer.

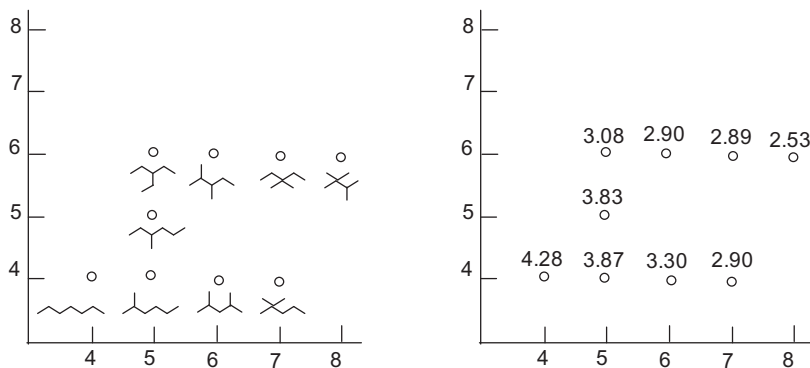


Fig. 3 Plot (p_2, p_3) for heptane isomers and their average matrix elements

Looking at Fig. 3 one observes that Q values (the average matrix element of D_{MAX} matrix for an isomer) decrease both with increase of the number of paths of length two and paths of length three. This clearly shows that there appears to be a structural regularity relating Q values of different isomers, and that this molecular descriptor may even be of some interest in structure-property-activity studies. We may characterize the dependence of Q values on (p_2, p_3) as $(-, -)$, and should we come across some molecular property that shows the same trends, decreasing values as the numbers of both p_2 and p_3 decrease, then Q molecular descriptors may be of interest. The Q values could also be of interest for correlations with molecular properties that show an increasing trend with the number of both p_2 and p_3 paths, properties that would be characterized as $(+, +)$, the difference in trends will be reflected in the sign of the regression line. In this respect it is of interest to recall that Randić and Wilkins [23] reported for octane isomers the critical density, the critical pressure and the molar magnetic susceptibilities as physicochemical properties having the $(+, +)$ trend. In addition, the Wiener number W in the case of octane isomers increases with the count of paths p_2 and p_3 . All this suggests that the average matrix element of D_{MAX} may be of interest in structure-property-activity studies as a novel molecular volume descriptor.

In Table 9 we have also listed sequences of ordered row sums, which are unique for each isomer. Row sums are of interest for the construction of Balaban's J index [24], a molecular topological index which has shown a fair discriminatory power, and has found use in quantitative structure-property relationship (QSPR) and quantitative structure-activity relationship (QSAR) studies. In view that calculations of row sums is so trivial the row sums may also be of potential interest in testing graphs for isomorphism.

6. D_{MAX} Invariants for Smaller Alkanes

We will end this introductory account of D_{MAX} matrices by Table 10 in which we have listed for smaller alkanes the number of non-zero entries in D_{MAX} of smaller alkanes; the sum of non-zero entries above the main diagonal; the quotient of the two, which gives the average D_{MAX} matrix element for each molecule; the partition of w coming from distances of decreasing length; and the ordered row sums.

Observe that except for the column Q , the average D_{MAX} matrix element for each molecule (and, of course, the last column of the Wiener numbers) the apparent “erratic” or chaotic variations of the remaining matrix invariants of Table 10. For example, the sum of the matrix elements of D_{MAX} above the main diagonal for pentane, isopentane and neopentane is: 14, 12, 16, while in the case of the full distance matrix it is: 20, 18, 16 (which gives the Wiener number W of these molecules). Even more dramatic are variations among the five isomers of

Table 10 The number of non-zero entries above the main diagonal in D_{MAX} of smaller alkanes (column #); the sum of non-zero entries above the main diagonal (column w); the quotient $w/\#$ giving the average matrix element (column Q); partition of w coming from distances of decreasing length (column π); ordered row sums (column Σ); and the Wiener number (column W).

	#	w	Q	π	Σ	W
ethane	1	1	1.0000	1	1, 1	1
Propane	3	4	1.3333	1, 2	3, 3, 2	4
butane	3	7	2.3333	1, 2, 2	5, 5, 2, 2	10
Isobutene	6	9	1.5000	0, 3, 3	5, 5, 5, 3	9
Pentane	5	14	2.8000	2, 2, 2, 0	9, 9, 4, 3, 3	20
isopentane	5	12	2.4000	0, 2, 3, 0	8, 5, 5, 4, 2	18
neopentane	10	16	1.6000	0, 0, 6, 4,	8, 8, 8, 8, 4	16
<i>n</i> -hexane	5	19	3.8000	1, 2, 2, 0, 0	12, 12, 4, 4, 3, 3	35
2-methylpentane	8	23	2.8750	0, 2, 3, 3, 0	11, 6, 5, 5, 5, 2	32
3-methylpentane	7	20	2.8571	0, 1, 4, 2, 0	12, 12, 6, 4, 3, 3	31

2,2-dimethylbutane	7	17	2.4286	0, 0, 3, 2, 0	11, 6, 5, 5, 5, 2	28
2,3-dimethylbutane	8	20	2.5000	0, 0, 4, 4, 0	8, 8, 8, 8, 4, 4	29

hexanes, which also show degeneracy, the w for 3-methylpentane and 2,3-dimethylbutane being equal because of having the same sum of non-zero elements above the main diagonal. This was the reason not to expect invariants of D_{MAX} to be of much interest in QSPR and QSAR studies, except in the special case, as is Q , the quotient of the sum of non-zero entries above the main diagonal and the number of non-zero entries above the main diagonal in D_{MAX} , which gives the average matrix element of D_{MAX} , and similar constructions.

7. Concluding Remarks

The D_{MAX} matrix of graphs here introduced may open new directions of exploration not only in Chemical Graph Theory but also possibly in Graph Theory on the role of dominant distances in graphs, as we may refer to the matrix elements of D_{MAX} matrix. A justification for this somewhat provocative statement is in the *generality* of the concept, which though illustrated here on distance matrix of graphs, allows one to extend its application to other graph matrices, and matrices in general. This includes a number of matrices mentioned in the introductory part of this article, like the resistance-distance matrix [8], the Szeged and revised Szeged matrices [5, 6], and the path Wiener matrix [4]. In addition one could consider the terminal matrix (for trees) [25-28], the line distance matrix [29, 30], and the pseudo-distance matrix of graphs obtained by considering columns of adjacency matrix as a vector in n -dimensional vector space [15].

Preliminary results on higher alkanes (octanes, nonanes and decanes) and several larger trees, indicates that the invariant w , which has shown degeneracy already in the case of hexane isomers, continues to show high degeneracy, even if somewhat less than the Wiener index W , which as the number of vertices increases tends to be excessive. Thus in this respect the use of the invariant w alone is not promising for isomorphism test. However, from D_{MAX} one can also construct an invariant, the construction of which is related to Balaban's J index [24], which is based on using row sums of the distance matrix to construct a topological index analogous to the connectivity index [31]. The J index of Balaban has been known as a topological index of high

discrimination power (relative to its simple conceptual and structural features), and the same, if not even better, appears to be the case with j index, constructed in a similar manner from the D_{MAX} matrix [32].

In this respect the ordered row sums (column Σ in Table 9 and 10) may be of considerable interest, in view of the ease with which they can be calculated, because they appear to show very limited degeneracy. Thus in case of the 18 octane isomers there is only one degenerate pair (3-ethylhexane and 2,3-dimethylhexane) but three pairs of isomers have duplicate w and three isomers have the same w , i. e. total of a nine octane isomers without unique w [32]. Similarly, in the case of the 35 nonane isomers there are three pairs of duplicate ordered row sum sequences among 22 non-unique w values. Hence, when one comes across graphs having the same Σ one should consider additional graph invariants, like the above mentioned indices J and j .

Acknowledgments

I would like to thank Dr. Marjana Novič, the head of the Laboratory for Chemometric, National Institute of Chemistry, Ljubljana, Slovenia, for warm hospitality and financial support. This work has been supported in part by the Ministry of Science and Higher Education, of Republic of Slovenia under Research Grant P1017. I would also like to thank Professor A. T. Balaban (Texas A&M University, Galveston, TX) and Professor Ivan Gutman (University of Kragujevac, Kragujevac, Serbia) for useful comments that improved the presentation of the material.

References

- [1] H. Poincaré, Second complément à l'Analysis Situs, *Proc. London Math. Soc.* **32** (1900) 277–308
- [2] F. Harary, *Graph Theory*, Addison-Wesley, Reading, 1969.
- [3] H. Wiener, Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **69** (1947) 17–20.
- [4] H. Wiener, Influence of interatomic forces on paraffin properties. *J. Chem. Phys.* **15** (1947) 766–766.

- [5] H. Wiener, Vapor pressure–temperature relationships among the branched paraffin hydrocarbons, *J. Phys. Chem.* **52** (1948) 425–430.
- [6] H. Wiener, Relation of the physical properties of the isomeric alkanes to molecular structure. Surface tension, specific dispersion, and critical solution temperature in aniline. *J. Phys. Chem.* **52** (1948) 1082–1089.
- [7] A. Graovac, T. Pisanski, On the Wiener index of a graph, *J. Math. Chem.* **8** (1991) 53–62.
- [8] M. Randić, Novel molecular descriptor for structure–property studies, *Chem. Phys. Lett.* **211** (1993) 478–483.
- [9] I. Gutman, A formula for the Wiener number of trees and its extension to graphs containing cycles, *Graph Theory Notes New York* **27** (1994) 9–15.
- [10] I. Gutman, S. Klavžar, An algorithm for the calculation of the Szeged index of benzenoid hydrocarbons, *J. Chem. Inf. Comput. Sci.* **35** (1995) 1011–1014.
- [11] T. Pisanski, M. Randić, Use of the Szeged index and the revised Szeged index for measuring network bipartivity, *Discr. Appl. Math.* **158** (2010) 1936–1944.
- [12] M. Randić, A. F. Kleiner, L. M. DeAlba, Distance/distance matrices, *J. Chem. Inf. Comput. Sci.* **34** (1994) 277–286.
- [13] D. J. Klein, M. Randić, Resistance distance, *J. Math. Chem.* **12** (1993) 81–95.
- [14] D. Babić, D. J. Klein, I. Lukovits, S. Nikolić, N. Trinajstić, Resistance–distance matrix: A computational algorithm and its applications, *Int. J. Quantum Chem.* **90** (2002) 166–176.
- [15] M. Randić, T. Pisanski, M. Novič, D. Plavšić, Novel graph distance matrix, *J. Comput. Chem.* **31** (2010) 1832–1841.
- [16] M. Randić, M. Plavšić, M. Razinger, Double invariants, *MATCH Commun. Math. Comput. Chem.* **35** (1997) 243–259.
- [17] M. Randić, (unpublished result)
- [18] D. Janežič, A. Miličević, S. Nikolić, N. Trinajstić, *Graph Theoretical Matrices in Chemistry*, Univ. Kragujevac, Kragujevac, 2007,
- [19] K. A. Zaretskii, Constructing a tree on the basis of a set of distances between the hanging vertices, *Uspekhi Mat. Nauk* **20** (1965) 90–92.

- [20] M. Randić, R. Orel, A. T. Balaban, D_{MAX} matrix for graphs having the same Balaban's J index, *MATCH Commun. Math. Comput. Chem.* **70** (2013) 239–258.
- [21] A. T. Balaban, A. Beteringhe, T. Constantinescu, P. A. Filip, O. Ivanciuc, Four new topological indices based on the molecular path code, *J. Chem. Inf. Model.* **47** (2007) 716–731.
- [22] M. Randić, C. L. Wilkins, On a graph theoretical basis for ordering of structures, *Chem. Phys. Lett.* **63** (1979) 332–336.
- [23] M. Randić, C. L. Wilkins, Graph theoretical ordering of structures as basis for systematic searches for regularities in molecular data, *J. Phys. Chem.* **83** (1979) 1525–1540.
- [24] A. T. Balaban, Highly discriminating distance-based topological index, *Chem. Phys. Lett.* **89** (1982) 399–404.
- [25] I. Gutman, B. Furtula, A survey on terminal Wiener index, in: I. Gutman, B. Furtula (Eds.), *Novel Molecular Structure Descriptors – Theory and Applications*, Univ. Kragujevac, Kragujevac, 2010, pp. 173–190.
- [26] I. Gutman, B. Furtula, M. Petrović, Terminal Wiener index, *J. Math. Chem.* **46** (2009) 522–531.
- [27] B. Horvat, T. Pisanski, M. Randić, Terminal polynomials and star-like graphs, *MATCH Commun. Math. Comput. Chem.* **60** (2008) 493–512.
- [28] M. Randić, J. Zupan, D. Vikić–Topić, Graphical representation of proteins by star-like graphs, *J. Mol. Graphics Model.* **26** (2007) 290–305.
- [29] M. Randić, J. Zupan, T. Pisanski, On representation of DNA by line distance matrix. *J. Math. Chem.* **43** (2008) 674–692.
- [30] G. Jaklič, T. Pisanski, M. Randić, On description of biological sequences by spectral properties of line distance matrix, *MATCH Commun. Math. Comput. Chem.* **58** (2007) 301–307.
- [31] M. Randić, Characterization of molecular branching, *J. Am. Chem. Soc.* **97** (1975) 6609–6615.
- [32] M. Randić (work in progress).