On the Harary index for the characterization of chemical graphs*

Dejan Plavšić, Sonja Nikolić, Nenad Trinajstić
The Rugjer Bošković Institute, P.O.B. 1016, 41001 Zagreb, Croatia

and

Zlatko Mihalić
Faculty of Science and Mathematics, The University of Zagreb, Strossmayerov trg 14, 41000 Zagreb, Croatia

"... mathematics to him ... was muscular strength and long walks over the hills and the kiss of a girl in love and big evenings spent swilling beer with your friends; ..."

John Dos Passos [1]

A novel topological index for the characterization of chemical graphs, derived from the reciprocal distance matrix and named the Harary index in honor of Professor Frank Harary, has been introduced. The Harary index is not a unique molecular descriptor; the smallest pair of the alkane trees with identical Harary indices has been detected in the octane family. The use of the Harary index in the quantitative structure–property relationships is exemplified in modeling physical properties of the C_2–C_9 alkanes. In this application, the performance of the Harary index is comparable to the performance of the Wiener number.

1. Introduction

The numerical invariants of chemical graphs are increasingly being used for a single number characterization of the corresponding chemical compounds [2]. These invariants are named in the chemical literature as topological indices [3] or graph-theoretical indices [4]. The former term is the more common of the two [5]. Topological indices have found application in various areas of chemistry, physics, mathematics, informatics, biology, etc. [2–12], but their most important use to date is in the non-empirical [13] quantitative structure–property relationships (QSPR) and quantitative structure–activity relationships (QSAR) [4,9,12,14–19].

*Dedicated to Professor Frank Harary on the happy occasion of his 70th birthday.

Here, we introduce a novel topological index for the characterization of chemical graphs which we name the Harary index in honor of Professor Frank Harary, the grandmaster of both graph theory and chemical graph theory. The Harary index, denoted by H, is derived from the reciprocal distance matrix and has a number of interesting properties. Its performance in QSPR is tested in the structure–property correlation with several physical properties of alkanes.

2. The definition of the Harary index and some of its properties

The Harary index $H = H(G)$ of a graph G is based on the concept of the reciprocal distance and can be obtained as a half-sum of the elements in the reciprocal distance matrix $\mathbf{D}^r = \mathbf{D}^r(G)$,

$$H = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\mathbf{D}^r)_{ij}. \tag{1}$$

The reciprocal distance matrix $\mathbf{D}^r$ is obtained by replacing all matrix elements $(\mathbf{D})_{ij}$, representing the shortest distances between vertices i and j, in the distance matrix $\mathbf{D} = \mathbf{D}(G)$ of a graph G by their reciprocals,

$$(\mathbf{D}^r)_{ij} = \frac{1}{(\mathbf{D})_{ij}}; \quad i \neq j. \tag{2}$$

The reciprocal distance matrix was also recently used by Ivanciuc et al. [20–22].

The Harary index for chains (depicting, for example, the carbon skeletons of the n-alkanes) is given by

$$H = N \sum_{k=1}^{N-1} \frac{1}{k} - (N-1), \tag{3}$$

where N is the number of sites (vertices) in the chain.

The Harary index for cycles (depicting, for example, the carbon skeletons of [N]annulenes) is given by

$$\begin{aligned} H &= N \sum_{k=1}^{\frac{1}{2}(N-2)} \frac{1}{k} + 1, \qquad N = \text{even}; \\ &= N \sum_{k=1}^{\frac{1}{2}(N-1)} \frac{1}{k}, \qquad N = \text{odd}. \end{aligned} \tag{4}$$

The above definition of the Harary index parallels the definition [3] of the Wiener number [23] $W = W(G)$ of a graph G,

$$W = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\mathbf{D})_{ij}, \tag{5}$$

where $(\mathbf{D})_{ij}$ represent the off-diagonal elements of the distance matrix **D** of G.

The Wiener numbers for chains and cycles can also be given in a closed form [24,25],

$$W = N(N^2 - 1)/6, \tag{6}$$

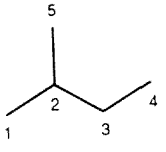
$$\begin{aligned} W &= (N/2)^3, && N = \text{even}; \\ &= N(N^2 - 1)/8, && N = \text{odd}. \end{aligned} \tag{7}$$

Both indices, the Harary index and the Wiener index, can be easily extended to unsaturated and heteroatomic systems through their distance matrices. There are several procedures available for the construction of the distance matrix for these systems (e.g. ref. [21]).

An example of calculating the reciprocal distance matrix and the Harary index for a tree T corresponding to 2-methylbutane is given in table 1.

Table 1

The distance matrix **D**, the reciprocal distance matrix $\mathbf{D}^r$, the Wiener number and the Harary index for a tree T depicting 2-methylbutane.



T

$$\mathbf{D} = \begin{bmatrix} 0 & 1 & 2 & 3 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 3 \\ 2 & 1 & 2 & 3 & 0 \end{bmatrix}$$

$W = 18$

$$\mathbf{D}^r = \begin{bmatrix} 0 & 1 & 0.5 & 0.33 & 0.5 \\ 1 & 0 & 1 & 0.5 & 1 \\ 0.5 & 1 & 0 & 1 & 0.5 \\ 0.33 & 0.5 & 1 & 0 & 0.33 \\ 0.5 & 1 & 0.5 & 0.33 & 0 \end{bmatrix}$$

$H = 6.66$

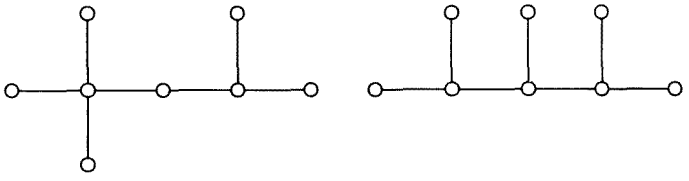
The reciprocal distance matrix $\mathbf{D}^r$ is introduced instead of the distance matrix **D** because the latter has an undesirable feature for applications in physics and chemistry, that is, the matrix elements of **D** related to the distant sites are associated with large entries. This is contrary to the intuitive expectation that the distant sites should influence each other much less than the near sites. Thus, the matrix elements related to the distant sites should be associated with smaller entries than the matrix elements corresponding to the vicinal sites.

A variant of the Harary index $H' = H'(G)$ of a graph G is based on the matrix containing as elements squares of reciprocal distances in G [26],

$$H' = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{(\mathbf{D}^2)_{ij}}. \tag{8}$$

The Harary index has several interesting properties. Some of these are listed below.

(1) The Harary index is a more discriminating index than the Wiener number, but it is not unique. The smallest pair of the alkane trees with the same index is given in fig. 1. The smallest pair of cyclic graphs with the same Harary index is given in fig. 2. This finding is obviously also valid for a variant of the Harary index H'.



T_1 T_2

$H(T_1) = 15.1667$ $H(T_2) = 15.1667$

Fig. 1. The smallest pair of alkane trees with the identical Harary indices.

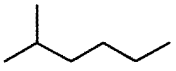
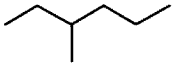
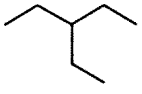
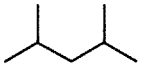
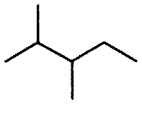
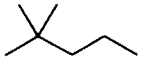
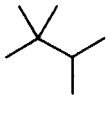
G_1 G_2

$H(G_1) = 11.3333$ $H(G_2) = 11.3333$

Fig. 2. The smallest pair of cyclic graphs with the identical Harary indices.

Table 2

The Harary indices (H) and the Wiener numbers (W) for heptane isomers.

Heptane isomer	H	W
	11.1500	56
	11.4833	52
	11.6167	50
	11.7500	48
	11.8333	48
	12.0000	46
	12.0833	46
	12.2500	44
	12.5000	42

(2) The Harary index appears to be a convenient measure of the compactness of the molecule. The larger the Harary index, the larger the compactness of the molecule. The inverse is true for the Wiener number. This is illustrated in table 2 for the family of heptanes. The least compact molecule appears to be n-heptane ($H = 11.15$, $W = 56$) and the most compact molecule 2,3,3-trimethylbutane ($H = 12.50$, $W = 42$).

The order of heptane isomers given in table 2 also follows the branching pattern of heptanes. Thus, the Harary index may also be used as a convenient measure of branching in alkane trees.

In compact structures, the number of short distances increases, hence the Harary index is larger than in the case of extended structures. Similarly, a more branched structure has more short distances and the Harary index is, thus, also larger for more branched structures.

The reciprocal distance matrix may also be used for detecting the center of a graph. The center of a graph G is the set of all central vertices [27]. A vertex v is a central graph if the maximum possible distance between v and any other vertex in G is as small as possible. This statement may be reformulated in terms of reciprocal distances as follows: A vertex v is a central vertex if the corresponding reciprocal distance sum d_v^r is larger than for any other vertex in G. The reciprocal distance sum is defined as

$$d_i^r = \sum_{j=1}^{N} (\mathbf{D}^r)_{ij}. \tag{9}$$

The application of the above procedure is exemplified for several graphs in fig. 3.

The above procedure agrees in most cases with the procedure by Bonchev et al. [28–31], which is based on the distance matrix and hierarchy of criteria. However, in a few cases we differ. For example, their procedure gives as the center of the graph G_4 (see fig. 3) vertex number 5.

3. The use of the Harary index in QSPR studies

We have used the Harary index and its variants for modeling eight representative physical properties (boiling points (bp), molar volumes (mv) at 20 °C, molar refractions (mr) at 20 °C, heats of vaporization (hv) at 25 °C, critical temperatures (ct), critical pressures (cp) and surface tensions (st) at 20 °C) of the 74 alkanes from ethane to nonanes. Values for these properties were taken from Needham et al. [32]. The Harary indices and the experimental values for the above physical properties of the 74 C_2–C_9 alkanes are listed in table 3. In this table are also given the corresponding Wiener numbers.

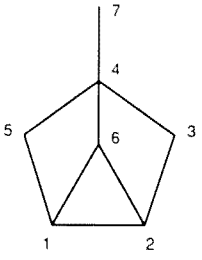
Three QSPR models were tested:

(a) linear model

$$P = A + B(\mathrm{TI}), \tag{10}$$

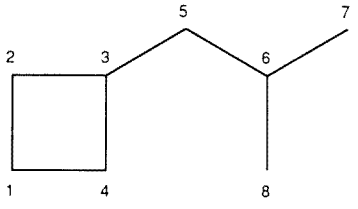
(b) quadratic model

$$P = A + B(\mathrm{TI}) + C(\mathrm{TI})^2, \tag{11}$$



G_3

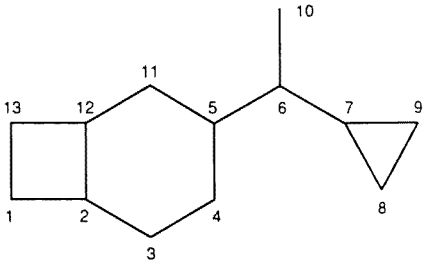
$$\mathbf{D}^r = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix} \\ \left[\begin{matrix} 0 & 1 & 0.5 & 0.5 & 1 & 1 & 0.33 \\ 1 & 0 & 1 & 0.5 & 0.5 & 1 & 0.33 \\ 0.5 & 1 & 0 & 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0.5 & 0.5 & 1 & 0 & 0.5 & 0.5 \\ 1 & 1 & 0.5 & 1 & 0.5 & 0 & 0.5 \\ 0.33 & 0.33 & 0.5 & 1 & 0.5 & 0.5 & 0 \end{matrix}\right] \end{array} \begin{array}{c} d_i^r \\ \\ 4.33 \\ 4.33 \\ 4.00 \\ 5.00 \\ 4.00 \\ 4.50 \\ 3.16 \end{array} \Leftarrow$$



G_4

$$\mathbf{D}^r = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \left[\begin{matrix} 0 & 1 & 0.5 & 1 & 0.33 & 0.25 & 0.2 & 0.2 \\ 1 & 0 & 1 & 0.5 & 0.5 & 0.33 & 0.25 & 0.25 \\ 0.5 & 1 & 0 & 1 & 1 & 0.5 & 0.33 & 0.33 \\ 1 & 0.5 & 1 & 0 & 0.5 & 0.33 & 0.25 & 0.25 \\ 0.33 & 0.5 & 1 & 0.5 & 0 & 1 & 0.5 & 0.5 \\ 0.25 & 0.33 & 0.5 & 0.33 & 1 & 0 & 1 & 1 \\ 0.2 & 0.25 & 0.33 & 0.25 & 0.5 & 1 & 0 & 0.5 \\ 0.2 & 0.25 & 0.33 & 0.25 & 0.5 & 1 & 0.5 & 0 \end{matrix}\right] \end{array} \begin{array}{c} d_i^r \\ \\ 2.98 \\ 3.83 \\ 4.66 \\ 2.83 \\ 4.33 \\ 4.41 \\ 3.03 \\ 3.03 \end{array} \Leftarrow$$

Fig. 3. Caption on following page.



G_5

$$\mathbf{D}^r = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \begin{matrix} & & & & & & & & & & & & & d_i^r \end{matrix} \\ \left[\begin{matrix} 0 & 1 & 0.5 & 0.33 & 0.25 & 0.2 & 0.17 & 0.14 & 0.14 & 0.17 & 0.33 & 0.5 & 1 \\ 1 & 0 & 1 & 0.5 & 0.33 & 0.25 & 0.2 & 0.17 & 0.17 & 0.2 & 0.5 & 1 & 5 \\ 0.5 & 1 & 0 & 1 & 0.5 & 0.33 & 0.25 & 0.2 & 0.2 & 0.25 & 0.33 & 0.5 & 0.33 \\ 0.33 & 0.5 & 1 & 0 & 1 & 0.5 & 0.33 & 0.25 & 0.25 & 0.33 & 0.5 & 0.33 & 0.25 \\ 0.25 & 0.33 & 0.5 & 1 & 0 & 1 & 0.5 & 0.33 & 0.33 & 0.5 & 1 & 0.5 & 0.33 \\ 0.2 & 0.25 & 0.33 & 0.5 & 1 & 0 & 1 & 0.5 & 0.5 & 1 & 0.5 & 0.33 & 0.25 \\ 0.17 & 0.2 & 0.25 & 0.33 & 0.5 & 1 & 0 & 1 & 1 & 0.5 & 0.33 & 0.25 & 0.2 \\ 0.14 & 0.17 & 0.2 & 0.25 & 0.33 & 0.5 & 1 & 0 & 1 & 0.33 & 0.25 & 0.2 & 0.17 \\ 0.14 & 0.17 & 0.2 & 0.25 & 0.33 & 0.5 & 1 & 1 & 0 & 0.33 & 0.25 & 0.2 & 0.17 \\ 0.17 & 0.2 & 0.25 & 0.33 & 0.5 & 1 & 0.5 & 0.33 & 0.33 & 0 & 0.33 & 0.25 & 0.2 \\ 0.33 & 0.5 & 0.33 & 0.5 & 1 & 0.5 & 0.33 & 0.25 & 0.25 & 0.33 & 0 & 1 & 0.5 \\ 0.5 & 1 & 0.5 & 0.33 & 0.5 & 0.33 & 0.25 & 0.2 & 0.2 & 0.25 & 1 & 0 & 1 \\ 1 & 0.5 & 0.33 & 0.25 & 0.33 & 0.25 & 0.2 & 0.17 & 0.17 & 0.2 & 0.5 & 1 & 0 \end{matrix}\right] \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{l} \\ 4.73 \\ 5.82 \\ 5.39 \\ 5.57 \\ 6.57 \Leftarrow \\ 6.36 \\ 5.73 \\ 4.54 \\ 4.54 \\ 4.39 \\ 5.82 \\ 6.06 \\ 4.90 \end{array}$$

Fig. 3. The identification of the center of several graphs. The arrow denotes the graph center.

(c) biparametric model

$$P = A(\mathrm{TI})^B N^C + Dp_3 + E, \tag{12}$$

where P = physical property, TI = topological index, N = the number of vertices, and p_3 = the number of paths of length 3. In the case of coefficients B and C being very small in the above model, we used the logarithmic model

$$P = A \ln(\mathrm{TI}) + B \ln(N) + Cp_3 + D. \tag{13}$$

The statistical parameters for the linear QSPR model are reported in table 4, for the quadratic QSPR model in table 5, and for the biparametric QSPR model in both versions in table 6.

Table 3

Harary indices in both versions (H and H'), Wiener numbers (W) and the experimental values for the eight physical properties of the 74 alkanes.

Alkane	H	H'	W	bp (°C)	mv (cm^3)	mr (cm^3)	hv (kJ)	ct (°C)	cp (atm)	st (dyne/cm)	mp (°C)
ethane	1.0000	1.0000	1	−88.630				32.27	48.2		−183.27
propane	2.5000	2.2500	4	−42.070				96.80	42.01		−187.69
butane	4.3333	3.6111	10	−0.500				152.01	37.47		−138.35
2-methylpropane	4.5000	3.7500	9	−11.730				134.98	36		−159.60
pentane	6.4167	5.0347	20	36.074	115.205	25.2656	26.42	196.62	33.31	16.00	−129.72
2-methylbutane	6.6667	5.2222	18	27.852	116.426	25.2923	24.59	187.70	32.9	15.00	−159.90
2,2-dimethylpropane	7.0000	5.5000	16	9.503	122.074	25.7243	21.78	160.60	31.57		−16.55
hexane	8.7000	6.4983	35	68.740	130.688	29.9066	31.55	234.70	29.92	18.42	−95.35
2-methylpentane	9.0000	6.7083	32	60.271	131.933	29.9459	29.86	224.90	29.95	17.38	−153.67
3-methylpentane	9.0833	6.7569	31	63.282	129.717	29.8016	30.27	231.20	30.83	18.12	−118.00
2,2-methylbutane	9.5000	7.0833	28	49.741	132.744	29.9347	27.69	216.20	30.67	16.30	−99.87
2,3-dimethylbutane	9.3333	6.9444	29	57.988	130.240	29.8104	29.12	227.10	30.99	17.37	−128.54
heptane	11.1500	7.9897	56	98.427	146.540	34.5504	36.55	267.55	27.01	20.26	−90.61
2-methylhexane	11.4833	8.2119	52	90.052	147.656	34.5908	34.80	257.90	27.2	19.29	−118.28
3-methylhexane	11.6167	8.2831	50	91.850	145.821	34.4597	35.08	262.40	28.1	19.79	−119.40
3-ethylpentane	11.7500	8.3542	48	93.475	143.517	34.2827	35.22	267.60	28.6	20.44	−118.60
2,2-dimethylpentane	12.0833	8.6319	46	79.197	148.695	34.6166	32.43	247.70	28.4	18.02	−123.81
2,3-dimethylpentane	12.0000	8.5417	46	89.784	144.153	34.3237	34.24	264.60	29.2	19.96	−119.10
2,4-dimethylpentane	11.8333	8.4444	48	80.500	148.949	34.6192	32.88	247.10	27.4	18.15	−119.24
3,3-dimethylpentane	12.2500	8.7292	44	86.064	144.530	34.3323	33.02	263.00	30	19.59	−134.46
2,3,3-trimethylbutane	12.5000	8.9167	42	80.882	145.191	34.3736	32.04	258.30	29.75	18.76	−24.91
octane	13.7429	9.5015	84	125.665	162.592	39.1922	41.48	296.20	24.64	21.76	−56.79
2-methylheptane	14.1000	9.7311	79	117.647	163.663	39.2316	39.68	288.00	24.8	20.60	−109.04

. . . continues

Table 3 (continued)

Alkane	H	H'	W	bp (°C)	mv (cm^3)	mr (cm^3)	hv (kJ)	ct (°C)	cp (atm)	st (dyne/cm)	mp (°C)
3-methylheptane	14.2667	9.8144	76	118.925	161.832	39.1001	39.83	292.00	25.6	21.17	– 120.50
4-methylheptane	14.3167	9.8369	75	117.709	162.105	39.1174	39.67	290.00	25.6	21.00	– 120.95
3-ethylhexane	14.4833	9.9203	72	118.534	160.072	38.9441	39.40	292.00	25.74	21.51	
2,2-dimethylhexane	14.7667	10.1756	71	106.840	164.285	39.2525	37.29	279.00	25.6	19.60	– 121.18
2,3-dimethylhexane	14.7333	10.1078	70	115.607	160.395	38.9808	38.79	293.00	26.6	20.99	
2,4-dimethylhexane	14.6500	10.0592	71	109.429	163.093	39.1300	37.76	282.00	25.8	20.05	– 137.50
2,5-dimethylhexane	14.4667	9.9656	74	109.103	164.697	39.2596	37.86	279.00	25	19.73	– 91.20
3,3-dimethylhexane	15.0333	10.3178	67	111.969	160.879	39.0087	37.93	290.84	27.2	20.63	– 126.10
3,4-dimethylhexane	14.8667	10.1789	68	117.725	158.814	38.8453	39.02	298.00	27.4	21.64	
3-ethyl-2-methylpentane	14.9167	10.2014	67	115.650	158.794	38.8362	38.52	295.00	27.4	21.52	– 114.96
3-ethyl-3-methylpentane	15.2500	10.4375	64	118.259	157.026	38.7171	37.99	305.00	28.9	21.99	– 90.87
2,2,3-trimethylpentane	15.4167	10.5764	63	109.841	159.526	38.9249	36.91	294.00	28.2	20.67	– 112.27
2,2,4-trimethylpentane	15.1667	10.4306	66	99.238	165.083	39.2617	35.13	271.15	25.5	18.77	– 107.38
2,3,3-trimethylpentane	15.5000	10.6250	62	114.760	157.292	38.7617	37.22	303.00	29	21.56	– 100.70
2,3,4-trimethylpentane	15.1667	10.3889	65	113.467	158.852	38.8681	37.61	295.00	27.6	21.14	– 109.21
2,2,3,3-trimethylbutane	16.0000	11.0000	58	106.470				270.80	24.5		
nonane	16.4607	11.0289	120	150.798	178.713	43.8423	46.44	322.00	22.74	22.92	– 53.52
2-methyloctane	16.8357	11.2633	114	143.260	179.773	43.8795	44.65	315.00	23.6	21.88	– 80.40
3-methyloctane	17.0262	11.3540	110	144.180	177.952	43.7296	44.75	318.00	23.7	22.34	– 107.60
4-methyloctane	17.1095	11.3887	108	142.480	178.150	43.7687	44.75	318.30	23.06	22.34	– 113.20
3-ethylheptane	17.3000	11.4794	104	143.000	176.410	43.6420	44.81	318.00	23.98	22.81	– 114.90
4-ethylheptane	17.3833	11.5142	102	141.200	175.685	43.4907	44.81	318.30	23.98	22.81	
2,2-dimethylheptane	17.5500	11.7225	104	132.690	180.507	43.9138	42.28	302.00	22.8	20.80	– 113.00
2,3-dimethylheptane	17.5500	11.6669	102	140.500	176.653	43.6369	43.79	315.00	23.79	22.34	– 116.00
2,4-dimethylheptane	17.5167	11.6408	102	133.500	179.120	4.37393	42.87	306.00	22.7	21.30	

2,5-dimethylheptane	17.4167	11.5958	104	136.000	179.371	43.8484	43.87	307.80	22.7	21.30	
2,6-dimethylheptane	17.2167	11.5003	108	135.210	180.914	43.9258	4.282	306.00	23.7	20.83	– 102.90
3,3-dimethylheptane	17.8833	11.8892	98	137.300	176.897	43.6870	42.66	314.00	24.19	22.01	
3,4-dimethylheptane	17.7667	11.7728	98	140.600	175.349	43.5473	43.84	322.70	24.77	22.80	
3,5-dimethylheptane	17.6333	11.7017	100	136.000	177.386	43.6378	42.98	312.30	23.59	21.77	
4,4-dimethylheptane	17.9833	11.9342	96	135.200	176.897	43.6022	42.66	317.80	24.18	22.01	
3-ethyl-2-methylhexane	17.8500	11.8075	96	138.000	175.445	43.6550	43.84	322.70	24.77	22.80	
4-ethyl-2-methylhexane	17.7167	11.7364	98	133.800	177.386	43.6472	42.98	330.30	25.56	21.77	
3-ethyl-3-methylhexane	18.2333	12.0661	92	140.600	173.077	43.2680	43.04	327.20	25.66	23.22	
3-ethyl-4-methylhexane	17.9833	11.8786	94	140.400	172.844	43.3746	43.95	312.30	23.59	23.27	
2,2,3-trimethylhexane	18.3500	12.1825	92	133.600	175.878	43.6226	41.91	318.10	25.07	21.86	
2,2,4-trimethylhexane	18.1833	12.0853	94	126.540	179.220	43.7638	40.57	301.00	23.39	20.51	– 120.00
2,2,5-trimethylhexane	17.9500	11.9692	98	124.084	181.346	43.9356	40.17	296.60	22.41	20.04	– 105.78
2,3,3-trimethylhexane	18.4833	12.2536	90	137.680	173.780	43.4347	42.23	326.10	25.56	22.41	– 116.80
2,3,4-trimethylhexane	18.2333	12.0661	92	139.000	173.498	43.4917	42.93	324.20	25.46	22.80	
2,3,5-trimethylhexane	17.9667	11.9239	96	131.340	177.656	43.6474	41.42	309.40	23.49	21.27	– 127.80
2,4,4-trimethylhexane	18.3167	12.1564	92	130.648	177.187	43.6598	40.84	309.10	23.79	21.17	– 133.38
3,3,4-trimethylhexane	18.6167	12.3247	88	140.460	172.055	43.3407	42.28	330.60	26.45	23.27	– 101.20
3,3-diethylpentane	18.5000	12.2083	88	146.168	170.185	43.1134	43.36	342.80	26.94	23.75	– 33.11
2,2-dimethyl-3-ethylpentane	18.5833	12.2986	88	133.830	174.537	43.4571	42.02	322.60	25.96	22.38	– 99.20
2,3-dimethyl-3-ethylpentane	18.7500	12.3958	86	142.000	170.093	42.9542	42.55	338.60	26.94	23.87	
2,4-dimethyl-3-ethylpentane	18.3333	12.1111	90	136.730	173.804	43.4037	42.93	324.20	25.46	22.80	– 122.20
2,2,3,3-tetramethylpentane	19.2500	12.7708	82	140.274	169.495	43.2147	41.00	334.50	27.04	23.38	– 9.90
2,2,3,4-tetramethylpentane	18.8333	12.4861	86	133.016	173.557	43.4359	41.00	319.60	25.66	21.98	– 121.09
2,2,4,4-tetramethylpentane	18.7500	12.4792	88	122.284	178.256	43.8747	38.10	301.60	24.58	20.37	– 66.54
2,3,3,4-tetramethylpentane	19.0000	12.5833	84	141.551	169.928	43.2016	41.75	334.50	26.85	23.31	– 102.12

Table 4

Statistical parameters for the linear QSPR models.

Physical property	N	A	B	r	s	F
			topological index = H			
bp	74	10.257 (±0.369)	− 43.34 (±5.61)	0.957	13.6	773
mv	69	4.846 (±0.176)	88.78 (±2.75)	0.959	4.9	760
mr	69	1.5104 (±0.0358)	16.463 (±0.559)	0.982	1.0	1780
hv	69	1.4308 (±0.0796)	16.97 (±1.24)	0.910	2.2	323
ct	74	12.677 (±0.457)	95.84 (±6.96)	0.956	16.9	768
cp	74	− 0.8759 (±0.0577)	39.847 (±0.878)	0.873	2.1	230
st	68	0.5100 (±0.0358)	13.100 (±0.562)	0.869	1.0	203
mp	56	3.280 (±0.921)	− 153.9 (±13.3)	0.436	31.4	13
			topological index = H'			
bp	74	17.047 (±0.561)	− 63.40 (±5.77)	0.963	12.5	923
mv	69	8.306 (±0.308)	76.53 (±3.25)	0.957	5.0	727
mr	69	2.5097 (±0.0640)	12.638 (±0.675)	0.980	1.0	1637
hv	69	2.441 (±0.141)	13.47 (±1.49)	0.904	2.3	300
ct	74	21.095 (±0.685)	70.78 (±7.05)	0.964)	15.3	948
cp	74	− 1.468 (±0.090)	41.688 (±0.929)	0.887	2.0	265
st	68	0.8730 (±0.0624)	11.823 (±0.662)	0.865	1.0	196
mp	56	5.55 (±1.50)	− 161.4 (±14.8)	0.450	31.2	14
			topological index = W			
bp	74	1.4256 (±0.0730)	4.25 (±5.66)	0.917	18.6	381
mv	69	0.6432 (±0.0190)	113.99 (±1.52)	0.972	4.1	1145
mr	69	0.19370 (±0.00676)	24.835 (±0.540)	0.962	1.5	822
hv	69	0.19843 (±0.00667)	23.766 (±0.534)	0.964	1.4	883
ct	74	1.697 (±0.106)	159.31 (±8.24)	0.883	27.1	255
cp	74	− 0.12693 (±0.00837)	36.155 (±0.649)	0.873	2.1	230
st	68	0.06203 (±0.00551)	16.190 (±0.444)	0.811	1.1	127
mp	56	0.406 (±0.140)	− 135.6 (±10.2)	0.367	32.5	8

From the statistical results, we learn:

(i) The biparametric model is found to be the most accurate in all cases.

(ii) The most accurate QSPR model for predicting the boiling points amongst considered models is the biparametric model (12) with the Wiener index.

(iii) The biparametric model with all three indices predicts accurately the molar volumes.

Table 5

Statistical parameters for the quadratic QSPR models.

Physical property	N	A	B	C	r	s	F
			topological index = H				
bp	74	−96.85 (±6.89)	21.30 (±1.21)	−0.4652 (±0.0500)	0.981	9.2	889
mv	69	59.64 (±8,38)	9.48 (±1.28)	−0.1704 (±0.0468)	0.966	4.5	456
mr	69	9.01 (±1.61)	2.697 (±0.246)	−0.04359 (±0.00895)	0.987	0.9	1203
hv	69	5.23 (±3.87)	3.299 (±0.593)	−0.0686 (±0.0216)	0.923	2.1	189
ct	74	31.06 (±8.79)	26.04 (±1.55)	−0.5631 (±0.0638)	0.979	11.7	833
cp	74	47.79 (±1.14)	−2.514 (±0.201)	0.06900 (±0.00829)	0.938	1.5	259
st	68	9.83 (±1.90)	1.020 (±0.285)	−0.0185 (±0.0103)	0.875	0.9	107
mp	56	−182.4 (±23.6)	9.31 (±4.24)	−0.259 (±0.178)	0.471	31.1	8
			topological index = H'				
bp	74	−115.42 (±8.47)	32.34 (±2.17)	−0.947 (±0.132)	0.979	9.6	813
mv	69	42.33 (±11.9)	16.05 (±2.63)	−0.414 (±0.139)	0.962	4.8	410
mr	69	4.25 (±2.41)	4.489 (±0.531)	−0.1015 (±0.0282)	0.983	1.0	971
hv	69	−1.37 (±5.59)	5.80 (±1.21)	−0.1797 (±0.0642)	0.915	2.2	169
ct	74	7.3 (±10.3)	39.77 (±2.65)	−1.156 (±0.161)	0.979	11.7	835
cp	74	50.41 (±1.32)	−4.032 (±0.338)	0.1586 (±0.0205)	0.940	1.5	271
st	68	7.88 (±2.60)	1.752 (±0.564)	−0.0465 (±0.0296)	0.870	1.0	101
mp	56	−194.2 (±27.6)	15.42 (±7.18)	−0.623 (±0.443)	0.481	30.9	8
			topological index = W				
bp	74	−44.60 (±4.56)	3.569 (±0.158)	−0.01743 (±0.00124)	0.979	9.6	805
mv	69	97.61 (±2.18)	1.2184 (±0.0685)	−0.004245 (±0.000496)	0.987	2.8	1226
mr	69	18.423 (±0.678)	0.4189 (±0.0213)	−0.001662 (±0.000154)	0.986	0.9	1179
hv	69	18.763 (±0.865)	0.3742 (±0.0271)	−0.001297 (±0.000196)	0.979	1.1	744
ct	74	90.51 (±7.20)	4.716 (±0.249)	−0.02455 (±0.00196)	0.965	15.2	480
cp	74	40.197 (±0.796)	−0.3043 (±0.0275)	0.002442 (±0.000217)	0.923	1.7	206
st	68	12.335 (±0.828)	0.1921 (±0.0254)	−0.000941 (±0.000180)	0.871	1.0	102
mp	56	−151.6 (±15.4)	1.137 (±0.546)	−0.00606 (±0.00438)	0.406	32.3	5

(iv) The same as in (ii) is also valid for predicting the molar refractions of alkanes.

(v) The most accurate QSPR model for predicting the heats of vaporization is the biparametric model with the Harary index.

(vi) The most accurate QSPR model for predicting the critical temperatures is the biparametric model with the Wiener index.

Table 6

Statistical parameters for the QSPR models (12) and (13).

Physical property	n	A	B	C	D	E	r	s	F
				topological index = H					
bp	74	96.87 (±6.74)	− 164.8 (±22.0)	5.752 (±0.342)	6.69 (±15.6)		0.9970	3.65	3892
mv	69	13.88 (±1.77)	0.1125 (±0.0229	0.9343 (±0.0526)	− 1.9394 (±0.0681)	42.97 (±4.70)	0.9994	0.640	12340
mr	69	4.550 (±0.178)	0.02384 (±0.00623)	0.9844 (±0.0160)	− 0.14060 (±0.00529)	2.385 (±0.436)	1.0000	0.0499	188500
hv	69	5.45 (±2.85)	− 1.280 (±0.216)	2.668 (±0.450)	0.7743 (±0.0320)	− 11.47 (±7.41)	0.9984	0.311	5055
ct	74	81.7 (±15.7)	310.7 (±28.2)	5.033 (±0.409)	− 170.5 (±17.5)		0.9944	6.20	2051
cp	74	8.39 (±1.81)	− 34.85 (±3.27)	0.5252 (±0.0474)	72.51 (±2.03)		0.9867	0.718	862
st	68	− 1402. (±349.)	0.649 (±0.242)	− 1.774 (±0.576)	6.168 (±62.4)	450.2 (±62.4)	0.9960	5.287	2122
				topological index = H'					
bp	74	− 134.99 (±8.67)	353.0 (±13.4)	2.364 (±0.226)	− 323.62 (±8.27)		0.9974	3.44	4397
mv	69	13.42 (±1.73)	0.1108 (±0.00236)	0.9704 (±0.0501)	− 1.9089 (±0.0654)	43.3 (±4.73)	0.9993	0.647	12080
mr	69	4.521 (±0.175)	0.02452(±0.00628)	0.9907 (±0.0150)	− 0.13933 (±0.00501)	2.409 (±0.434)	1.0000	0.0497	190000
hv	69	− 55.65 (±2.40)	102.50 (±2.93)	0.5348 (±0.0673)	− 49.29 (±1.32)		0.9974	0.435	2502
ct	74	− 80.7 (±15.7)	287.6 (±24.2)	4.899 (±0.410)	− 154.6 (±15.0)		0.9943	6.23	2032
cp	74	− 57.1 (±39.4)	− 0.0435 (±0.0213)	0.300 (±0.115)	0.6211 (±0.0664)	118.3 (±43.3)	0.9874	0.706	669
st	68	0.031 (±0.079)	− 5.85 (±2.53)	8.82 (±3.79)	0.7082 (±0.0358)	12.05 (±1.72)	0.9808	0.388	399
				topological index = W					
bp	74	300.9 (±22.7)	0.2340 (±0.0144)	− 0.3280 (±0.0291)	4.847 (±0.165)	− 330.1 (±21.1)	0.9994	1.59	15390
mv	69	13.02 (±1.70)	− 0.04938 (±0.00913)	1.2269 (±0.0528)	− 1.9866 (±0.0715)	39.01 (±5.00)	0.9994	0.628	12810
mr	69	4.496 (±0.178)	0.00982 (±0.00266)	1.0445 (±0.0148)	− 0.14206 (±0.00571)	2.131 (±0.453)	1.0000	0.050	185500
hv	69	16.84 (±3.09)	0.954 (±0.118)	− 1.762 (±0.237)	0.8967 (±0.0426)	7.52 (±2.79)	0.9975	0.389	3223
ct	74	86.95 (±8.61)	− 115.9 (±28.0)	7.972 (±0.436)	107.8 (±20.0)		0.9968	4.66	3646
cp	74	− 36.0 (±11.5)	0.0297 (±0.0203)	0.2502 (±0.0856)	0.5947 (±0.0750)	91.0 (±12.4)	0.9874	0.707	673
st	68	37.7 (±16.2)	1.413 (±0.821)	− 3.70 (±2.15)	0.7832 (±0.0342)	8.49 (±4.23)	0.9862	0.330	557

(vii) A more accurate QSPR model for predicting the critial pressures is the biparametric model with the Wiener index and the modified Harary index.

(viii) The most accurate QSPR model for predicting the surface tensions is the biparametric model with the Harary index.

(ix) All models are rather inaccurate for predicting melting points. This result was also reached by many other authors (e.g. ref. [32]).

4. Concluding remarks

A novel topological index, named the Harary index in honor of Professor Frank Harary, has been introduced. It is derived from the reciprocal distance matrix. The Harary index has many interesting properties. Among these is the property of fair discriminating power, but it is not a unique molecular descriptor. Pairs of graphs with identical values of the Harary index have been detected. The Harary index was also tested in the QSPR modeling of physical properties of the alkanes. In this application, it is comparable to the Wiener number, as has been conjectured, since these two indices are intercorrelated quantities.

Acknowledgement

The financial support by the Ministry for Science, Technology and Informatics of the Republic of Croatia is gratefully acknowledged. We are thankful to Professor Frank Harary (Las Cruces) for helpful suggestions.

References

[1] J. Dos Passos, *The 42nd Parallel* (The Modern Library, New York, 1937) p. 325.
[2] N. Trinajstić, *Chemical Graph Theory*, 2nd. rev. Ed. (CRC, Boca Raton, 1992) ch. 10.
[3] H. Hosoya, Bull. Chem. Soc. Japan 44(1971)2332.
[4] S.C. Basak, G.J. Niemi and G.D. Veith, in: *Computational Chemical Graph Theory*, ed. D.H. Rouvray (Nova Science, Commack, New York, 1990) p. 235.
[5] D.H. Rouvray, Sci. Am. 254(1986)40.
[6] S.V. Nizhnii and N.A. Epshtein, Russian Chem. Rev. 47(1978)383.
[7] D. Bonchev, *Information Theoretic Indices for Characterization of Chemical Structures* (Research Studies Press, Chichester, 1983).
[8] H. Hosoya, in: *Mathematics and Computational Concepts in Chemistry*, ed. N. Trinajstić (Horwood, Chichester, 1986) p. 110.
[9] S.C. Basak, L.J. Monsrud, M.E. Rosen, C.M. Frane and V.R. Magnuson, Acta Pharm. Jugosl. 36(1986)81.
[10] E. Hladká and L. Matyska, Chemické Listy 82(1988)1009.
[11] M.I. Stankevitch, I.V. Stankevich and N.S. Zefirov, Russian Chem. Rev. 57(1988)191.
[12] N. Trinajstić, S. Nikolić and S. Carter, Kem. Ind (Zagreb) 38(1989)469.
[13] N. Trinajstić, M. Randić and D.J. Klein, Acta Pharm. Jugosl. 36(1986)267.
[14] L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research* (Academic Press, New York, 1976).

[15] A. Sabljić and N. Trinajstić, Acta Pharm. Jugosl. 31(1981)71.
[16] L.B. Kier and L.H. Hall, *Molecular Connectivity in Structure–Activity Analysis* (Research Studies Press, Letchworth, 1986).
[17] P.G. Seybold, M. May and U.A. Bagal, J. Chem. Educ. 64(1987)575.
[18] A. Sabljić, Environ. Health Persepctives 83(1989)179.
[19] A. Sabljić, in: *Practical Applications of Quantitative Structure–Activity Relationships (QSAR) in Environmental Chemistry and Toxicology*, ed. W. Karcher and J. Devillers (Kluwer, Dordrecht, 1990) p. 61.
[20] O. Ivanciuc, Rev. Roum. Chim. 34(1989)1361.
[21] T.-S. Balaban, P.A. Filip and O. Ivanciuc, J. Math. Chem. 11(1992)79.
[22] O. Ivanciuc, T.-S. Balaban and A.T. Balaban, J. Chem. Inf. Comput. Sci., submitted.
[23] H. Wiener, J. Am. Chem. Soc. 69(1947)17.
[24] D. Bonchev and N. Trinajstić, J. Chem. Phys. 67(1977)4517.
[25] O.E. Polansky, in: *MATH/CHEM/COMP/1988*, ed. A. Graovac (Elsevier, Amsterdam, 1989) p. 167.
[26] Z. Mihalić and N. Trinajstić, J. Chem. Educ. 69(1992)701.
[27] F. Harary, *Graph Theory*, 2nd printing (Addison–Wesley, Reading, 1971).
[28] D. Bonchev, A.T. Balaban and O. Mekenyan, J. Chem. Inf. Comput. Sci. 20(1980)106.
[29] D. Bonchev, A.T. Balaban and M. Randić, Int. J. Quant. Chem. 19(1981)61; erratum Int. J. Quant. Chem. 22(1982)441.
[30] D. Bonchev, O. Mekenyan and A.T. Balaban, J. Chem. Inf. Comput. Sci. 29(1989)91.
[31] D. Bonchev, J. Mol. Struct. THEOCHEM 185(1989)155.
[32] D.E. Needham, I.-C. Wei and P.G. Seybold, J. Am. Chem. Soc. 110(1988)4186.