Topological Indices and Real Number Vertex Invariants Based on Graph Eigenvalues or Eigenvectors

ALEXANDRU T. BALABAN,*,† DAN CIUBOTARIU,‡ and MIHAI MEDELEANU‡

Organic Chemistry Department, Polytechnic Institute, Splaiul Independentei 313, 77206 Bucharest, Romania, and Organic Chemistry Department, Polytechnic Institute, Timişoara, Romania

Eigenvectors obtained from the adjacency or distance matrix of graphs, corresponding to the largest negative eigenvalue, are useful real number vertex invariants for intramolecular ordering of vertices. From them or from eigenvalues, several new topological indices have been tested on the basis of intermolecular ordering of isomeric alkanes and of correlating ability with properties such as the boiling temperature.

INTRODUCTION

In the search for new graph invariants, real number vertex invariants (as opposed to integers) face better prospects because they may have lower degeneracy. A recent review on real number invariants[1] revealed that the oldest such invariants, namely, eigenvalues derived from the adjacency (A) or distance (D) matrices, deserve to be reinvestigated systematically.

As indicated in an older review,[2] topological indices (TI's) based on A or D can be derived either via primitive (or bookkeeping) operations or via sophisticated (or structural) operations: the former ones do not depend on the adjacencies of the vertex being examined and take into account only local properties of this vertex; the latter operations depend on longer range properties, being based on edges, paths, or the whole graph.

From the adjacency matrix A, eigenvalues are obtained via such a structural operation. Lovász and Pelikàn investigated eigenvalues of trees and found algebraic formulas for special types of trees such as path or star, and for particular eigenvalues such as the lowest eigenvalue denoted by $E(A)$.[3] No use of eigenvectors has apparently been made so far for correlations (see, however, below).

In this paper we examine eigenvalues and eigenvectors of all alkanes with 2–9 carbon atoms. In order to connect graph invariants with molecular properties, we shall present correlations with boiling points at normal pressure. It should be clear from the outset, however, that one is not to expect high-correlation coefficients because the boiling points of alkanes (which decrease with branching in each series of isomers) are lowered considerably by branching at the penultimate carbon atoms (i.e., linear chains flanked on both sides by two *tert*-butyl groups have the lowest boiling points among all isomeric alkanes); this feature is not necessarily reflected in the TI's.

We consider that a more important criterion for good local vertex invariants (LOVI's) is the intramolecular ordering of vertices, whereas for global (molecular) TI's the best criterion is the intermolecular ordering of alkanes among isomers. In both cases, intuitive or objective criteria about branching according to Bonchev and Trinajstić,[4] Gutman,[5] Randić,[5-7] Bertz,[8] and Balaban[2,9] have to be fulfilled. This ordering of alkanes was discussed in connection with the topological index[2,9] denoted by J.

In connection with the intramolecular ordering of vertices, mention should be made that Randić proposed to use the first eigenvector for the purpose of vertex labeling.[10]

This paper considers only acyclic systems; it may be extended to cyclic molecules, but not to regular graphs with high

†Polytechnic Institute, Bucharest.
‡Polytechnic Institute, Timişoara.

symmetry. Eigenvalues of alkanes have recently been used as topological indices.[11]

NEW LOCAL VERTEX INVARIANTS (LOVI'S)

From the adjacency (A) or distance matrices (D) of alkanes, by means of a computer program, all eigenvalues and eigenvectors were calculated for alkanes with 2–9 carbon atoms.

The eigenvectors corresponding to the lowest (largest negative) eigenvalue obtained from the adjacency matrix, $E(A)$, vary in a consistent manner and induce a reasonable intramolecular ordering of vertices (unlike eigenvectors corresponding to the other eigenvalues). Table I presents these eigenvectors for isomeric heptanes and octanes with the IUPAC convention for the vertex numbering. The longest chain is numbered sequentially starting from one end, and the side-chain carbons are then numbered further beginning with the atoms closest to the carbon 1 of the longest chain. In all tables, alkane names are indicated by M (methyl) or E (ethyl) attached to C_m where m is the number of carbon atoms in the longest chain.

It will be observed that the *lower values* of the eigenvectors correspond to vertices of lower degree, farther from the center, or from a vertex of high degree. It may also be seen that chemically nonequivalent vertices may have equal (accidentally degenerate) LOVI's indicated in italics; this is the case of 2,4-dimethylpentane and 2,5-dimethylhexane, whose vertices of degree three and the central vertices of degree two have the same LOVI. 2-Methylhexane, 4-methylheptane, and 2,3-dimethylhexane also have identical LOVI's for the (first) methyl side chain and for the methylene carbon farthest from the side chain.

On comparing the eigenvectors corresponding to the unique negative eigenvalue obtained from the distance matrix which is denoted by $E(D)$, it can be seen (Table II) that the intramolecular vertex ordering obeys the same rules as above with the difference that the italicized words "*lower values*" are to be replaced by "*higher values*". No accidental degeneracies of eigenvectors have been observed for isomers from Table II.

Eigenvectors derived from other eigenvalues corresponding to D do not lead to reasonable LOVI's.

One may conclude that eigenvectors corresponding to the largest negative eigenvalue $E(A)$ or $E(D)$ can serve as real number LOVI's. Those derived from D have lower degeneracy than those derived from A.

NEW TOPOLOGICAL INDICES

From the eigenvalues and eigenvectors computed for all C_2–C_9 alkanes, we calculated the following six types of TI's, all bearing the final number 1 (TI's of type 1):

VAA1 is the sum of positive eigenvalues from A.

Table I. Eigenvectors (LOVI's) and Eigenvalues Based on the Adjacency Matrix[a]

alkane	carbon atom number[b]								$-E(A)$[c]
	1	2	3	4	5	6	7	8	
C7	1913	3536	4619	5000	4619	3536	1913		1.848
2M-C6	*2887*	5577	5000	4083	*2887*	1494	*2887*		1.932
3M-C6	2073	4083	5968	4642	3176	1612	3030		1.970
24MM-C5	2500	*5000*	*5000*	*5000*	2500	2500	2500		*2.000*
3E-C5	2041	4083	6124	4083	2041	4083	2041		*2.000*
23MM-C5	2603	5344	5765	3682	1793	2603	2808		2.053
22MM-C5	3109	6533	4397	2706	1288	3109	3109		2.101
33MM-C5	1845	3941	6572	3941	1845	3077	3077		2.136
223MMM-C4	2887	6280	5000	2299	2887	2887	2299		2.175
C8	1612	3030	4083	4642	4642	4083	3030	1612	1.879
2M-C7	2673	5211	4816	4179	3333	2319	1189	2673	1.950
3M-C7	1927	3834	5698	4635	3522	2369	1191	2865	1.989
4M-C7	1443	*2887*	4330	5774	4330	*2887*	1443	*2887*	*2.000*
25MM-C6	2236	*4472*	*4472*	*4472*	*4472*	2236	2236	2236	*2.000*
3E-C6	1899	3851	5914	4294	2796	1378	3851	1899	2.029
24MM-C6	2188	4467	4747	5227	3367	1649	2188	2560	2.042
23MM-C6	*2454*	5090	5651	3907	*2454*	1183	*2454*	2724	2.074
34MM-C6	1601	3355	5429	5429	3355	1601	2591	2591	2.095
3E-2M-C5	2380	5000	5745	3536	1683	2380	3536	1683	2.101
22MM-C6	3019	6376	4409	2936	1792	0849	3019	3019	2.112
234MMM-C5	2176	4647	5574	4647	2176	2176	2610	2176	2.136
224MMM-C5	2799	6015	4529	3718	1730	2799	2799	1730	2.149
33MM-C6	1755	3786	6409	4093	2418	1121	2972	2972	2.157
3E-3M-C5	1696	3713	6432	3713	1696	2938	3713	1696	2.189
223MMM-C5	2729	6019	5092	2905	1317	2729	2729	2308	2.206
233MMM-C5	2115	4700	6210	3505	1578	2115	2795	2795	2.222
2233MMMM-C4	2454	5651	5651	2454	2454	2454	2454	2454	2.303

[a] Eigenvectors corresponding to the lowest eigenvalue (indicated here with changed sign). The eigenvectors are multiplied with 10 000 and presented with four significant digits. Accidentally degenerate values are in italics, when nonequivalent vertices share the same values. Standard numbering of carbon atoms in the main chain has been used, with side chains having subsequent numbers. [b] Eigenvectors. [c] Eigenvalues.

Table II. Eigenvectors (LOVI's) and Eigenvalues Based on the Distance Matrix[a]

alkane	carbon atom number[b]								$-E(D)$[c]
	1	2	3	4	5	6	7	8	
C7	4698	3697	3141	2963	3141	3697	4698		16.625
2M-C6	4240	3099	2911	3101	3694	4766	4240		15.405
3M-C6	4561	3425	2750	2978	3607	4722	3966		14.864
3E-C5	4494	3305	2579	3305	4494	3305	4494		14.297
24MM-C5	4282	3047	2846	3047	4282	4282	4282		14.176
23MM-C5	4208	2917	2670	3401	4630	4208	3993		13.635
22MM-C5	4005	2683	2929	3604	4809	4005	4005		13.635
33MM-C5	4575	3286	2499	3286	4575	3893	3893		13.070
223MMM-C4	4035	2583	2850	4265	4035	4035	4265		12.394
C8	4460	3594	3058	2801	2801	3058	3594	4460	21.836
2M-C7	4024	3057	2782	2778	3046	3611	4528	4024	20.479
3M-C7	4296	3324	2688	2700	2986	3574	4523	3719	19.763
4M-C7	4451	3484	2874	2559	2874	3484	4450	3611	19.542
25MM-C6	4073	3041	2753	2753	3041	4073	4073	4073	19.111
3E-C6	4180	3147	2449	2792	3432	4438	3147	4180	18.779
22MM-C6	3772	2670	2677	2975	3596	4608	3772	3772	18.413
24MM-C6	4050	2977	2668	2649	3326	4365	4050	3754	18.396
33MM-C6	4077	3009	2270	2653	3327	4363	4002	4002	18.308
23MM-C6	3957	2862	2517	2851	3498	4529	3957	3646	18.181
34MM-C6	4352	3272	2562	2562	3272	4352	3714	3714	17.676
3E-2M-C5	3918	2773	2396	3142	4250	3918	3142	4250	17.419
224MMM-C5	3811	2622	2636	2959	4113	3811	3811	4113	17.034
234MMM-C5	4010	2831	2465	2831	4010	4010	3684	4010	16.808
3E-3M-C5	4236	3079	2291	3079	4236	3533	3079	4236	16.670
223MMM-C5	3771	2527	2518	3278	4440	3771	3771	3763	16.315
233MMM-C5	3985	2749	2352	3151	4342	3985	3631	3631	16.068
2233MMMM-C4	3825	2470	2470	3825	3825	3825	3825	3825	14.937

[a] Eigenvectors corresponding to the lowest eigenvalue (indicated here with changed sign). The eigenvectors are multiplied with 10 000 and presented with four significant digits. [b] Eigenvectors. [c] Eigenvalues.

VAD1 is the negative eigenvalue $E(D)$.

VEA1 is the sum of elements (LOVI's) of the first eigenvector from A (cf. Table I).

VED1 is the sum of elements of the first eigenvector derived from the above eigenvalue $E(D)$, cf. Table II.

VRA1 is calculated from the LOVI's based on A by a Randić-type formula.

VRD1 is calculated likewise from LOVI's based on D.

The Randić-type formula is a structural operation:

$$VRA1 = \sum(a_i a_j)^{-1/2}$$

where a_i and a_j are LOVI's for endpoints of each edge, and the summation is extended over all graph edges.

Table III. Data for Alkanes with 2–8 Carbon Atoms

alkane	BP (°C)	VAA1	VAA2	VAA3	VAD1	VAD2	VAD3
C2	−88.5	1.00000	0.50000	−1.60944	1.00000	0.50000	−1.60944
C3	−44.5	1.41421	0.47140	−0.85740	2.73205	0.91068	−0.19892
C4	−0.5	2.23607	0.55902	−0.11157	5.16228	1.29057	0.72509
2M-C3	−10.5	1.73205	0.43301	−0.36698	4.64575	1.16144	0.61966
C5	36.5	2.73205	0.54641	0.31191	8.28822	1.65764	1.42169
2M-C4	27.9	2.61313	0.52263	0.26740	7.45929	1.49186	1.31631
22MM-C3	9.5	2.00000	0.40000	0.00000	6.60555	1.32111	1.19476
C6	68.7	3.49396	0.58233	0.74021	12.10931	2.01822	1.98315
3M-C5	63.2	3.44949	0.57941	0.72740	10.74240	1.79040	1.86337
2M-C5	60.2	3.07768	0.51295	0.61335	11.05878	1.84313	1.89240
23MM-C4	58.1	3.00000	0.50000	0.58779	10.00000	1.66667	1.79176
22MM-C4	49.7	2.90931	0.48489	0.55709	9.67021	1.61170	1.75822
C7	98.4	4.02734	0.57533	1.03643	16.62538	2.37505	2.45426
3E-C5	93.5	4.00000	0.57143	1.02962	14.29695	2.04242	2.30337
3M-C6	91.8	3.93923	0.56275	1.01431	14.86358	2.12337	2.34224
2M-C6	90.0	3.86370	0.55196	0.99495	15.40477	2.20068	2.37800
23MM-C5	89.8	3.83149	0.54736	0.98658	13.63462	1.94780	2.25594
33MM-C5	86.0	3.79793	0.54256	0.97778	13.06981	1.86712	2.21363
24MM-C5	80.5	3.41421	0.48774	0.87127	14.17597	2.02514	2.29487
22MM-C5	79.2	3.36028	0.48004	0.85535	13.63526	1.94789	2.25598
223MMM-C4	80.9	3.30136	0.47162	0.83766	12.39448	1.77064	2.16058
C8	125.8	4.75887	0.59485	1.44685	21.83635	2.78244	2.86043
3M-C7	118.8	4.70463	0.58808	1.32540	19.76280	2.47035	2.76066
3E-C6	118.9	4.72294	0.59037	1.32929	18.77877	2.34735	2.70958
34MM-C6	118.7	4.66587	0.58323	1.31713	17.67587	2.20498	2.64906
3E-3M-C5	118.2	4.64575	0.58072	1.31281	16.67049	2.08381	2.59050
4M-C7	117.7	4.41421	0.55178	1.26169	19.54202	2.44275	2.74942
2M-C7	117.6	4.38129	0.54766	1.25420	20.47922	2.55990	2.79627
3E-2M-C5	115.6	4.36028	0.54504	1.24939	17.41874	2.17734	2.63440
23MM-C6	115.3	4.32353	0.54044	1.24093	18.18148	2.27269	2.67726
24MM-C6	109.4	4.28259	0.53532	1.23142	18.39644	2.29956	2.68901
33MM-C6	112.0	4.25963	0.53245	1.22604	18.37416	2.29677	2.68780
25MM-C6	108.4	4.23607	0.52951	1.22049	19.64233	2.38894	2.75454
234MMM-C5	113.4	4.21215	0.52652	1.21483	16.80790	2.10099	2.59871
233MMM-C5	114.6	4.18757	0.52345	1.20898	16.06828	2.00853	2.55370
22MM-C6	107.0	4.15642	0.51955	1.20151	18.41326	2.30166	2.68993
223MMM-C5	110.5	4.13056	0.51632	1.19527	16.31517	2.03940	2.56895
224MMM-C5	99.3	3.69232	0.46154	1.08311	17.03382	2.12923	2.61206
2233MMMM-C4	106.0	3.60555	0.45069	1.05933	14.93725	1.86716	2.48071

Two kinds of normalization against the number n of carbon atoms of the alkane were carried out. Each of these led to a type of TI distinguished by the final number 2 or 3 as follows for the above TI's:

(i) for VAA1 and VAD1 (which increase progressively with increasing n), division by n, converting TI1 into TI2

(ii) multiplication by $n/10$ and natural logarithmation for converting TI1 into TI3

Thus we obtain new TI's that will be discussed in the following pages: TI2 = TI1/n; TI3 = ln (nTI1/10).

INTERMOLECULAR ORDERING OF ALKANES

In Tables 1 and 2, alkanes with 7 and 8 carbon atoms had been ordered according to the increasing eigenvalues $E(A)$ and to decreasing eigenvalues $E(D)$. On comparing the intermolecular ordering of heptane and octane isomers, it may be seen that there exist some differences between the order in the above two tables. For heptane isomers, the ordering based on eigenvalues of A is identical to the "ideal" one based on the J index[2,9] or on the graph derivatives.[8] The ordering based on eigenvalues of D differs only by the permutation 3-ethylpentane/2,4-dimethylpentane.

For octane isomers, the ordering based on A is again similar to that according to Bertz's graph derivatives, with the difference that the latter ordering places 2,5-dimethylhexane as being less branched than on the basis of eigenvalues of A. In the comparison with J, dimethylhexane isomers are ordered

by J as follows in increasing branching order: 2,5-; 2,4-, 2,2-; 2,3-; 3,4-; 3,3-, while the graph derivatives or the eigenvalues of A give the ordering: 2,5-; 2,4-; 2,3-; 3,4-; 2,2-; 3,3-. One can see that 2,2-dimethylhexane is "less branched" according to J. The ordering based on eigenvalues of D for octane isomers places dimethylhexanes in the following sequence: 2,5-; 2,2-; 2,4-; 3,3-; 2,3-; and 3,4-, hence quaternary carbon atoms are assigned much less "branching weight" than in the case of eigenvalues of A. The differences in ordering based on eigenvalues of D from the ordering based on J or the graph derivatives are much more pronounced than for the analogous comparison based on eigenvalues of A.

Tables III–VI present the new TI's (VAA, VAD, VEA, VED, VRA, and VRD) in the three versions each (1–3), separately for the alkanes with 2–8 carbon atoms, and 9 carbon atoms, respectively. Alkanes have been ordered according to decreasing VAA values because, despite the degeneracy of this index, it presents fairly similar characteristics to the variation of boiling points. Alkanes with side chains at the middle of the chain appear as less branched than those with side chains attached close to the ends of the chain.

Up to eight carbon atoms no degeneracy in the TI values was found. However, for nine carbon atoms, just one pair of isomers for VED indices and six pairs of isomers for VAA indices were found to have degenerate values (italicized in Tables V and VI).

Although all TI's described above have been computed and tested, only a few of these will be presented here in detail because few display a "rational intermolecular ordering" of

Table IV. Data for Alkanes with 2–8 Carbon Atoms

alkane	VEA1	VEA3	VED1	VED3	VRA1	VRD1
C2	1.41421	−1.26286	1.41421	−1.26286	1.41421	1.41421
C3	1.70711	−0.66917	1.71563	−0.66419	3.36358	3.72243
C4	1.94650	−0.25026	1.97417	−0.23614	5.89199	6.52546
2M-C3	1.93185	−0.25781	1.97226	−0.23711	5.58360	6.90091
C5	2.15470	0.07450	2.20361	0.09695	8.98667	9.73947
2M-C4	2.13099	0.06344	2.20196	0.09620	8.62989	10.15828
22MM-C3	2.12132	0.05889	2.20397	0.09711	8.00002	10.74143
C6	2.34190	0.34014	2.41180	0.36955	12.62883	13.31649
3M-C5	2.30080	0.32243	2.40851	0.36818	12.26119	13.88003
2M-C5	2.31281	0.32764	2.41168	0.36950	12.39112	13.67976
23MM-C4	2.30940	0.32616	2.41209	0.36967	11.52993	14.14868
22MM-C4	2.28550	0.31576	2.41110	0.36926	11.63736	14.40727
C7	2.51367	0.56507	2.60364	0.60024	16.80002	17.22301
3E-C5	2.44949	0.53920	2.59754	0.59789	16.39230	18.05924
3M-C6	2.45839	0.54283	2.60089	0.59918	16.73877	17.78548
2M-C6	2.48138	0.55214	2.60499	0.60075	16.81935	17.51358
23MM-C5	2.45983	0.54342	2.60267	0.59986	15.71166	18.19395
33MM-C5	2.42986	0.53116	2.60050	0.59903	15.79398	18.55190
24MM-C5	2.50000	0.55962	2.60668	0.60140	15.31371	17.86571
22MM-C5	2.42522	0.52925	2.60382	0.60031	16.77771	18.20072
223MMM-C4	2.45369	0.54092	2.60656	0.60136	14.73031	15.84788
C8	2.67347	0.76023	2.78244	0.80018	21.48306	21.43353
3M-C7	2.60405	0.73392	2.78102	0.79967	22.12919	21.93652
3E-C6	2.58816	0.72780	2.77621	0.79794	21.55154	22.33865
34MM-C6	2.59508	0.73047	2.77988	0.79926	20.49061	22.59672
3E-3M-C5	2.55992	0.71683	2.77683	0.79817	20.39206	23.11878
4M-C7	2.59808	0.73163	2.77872	0.79885	21.90424	22.06447
2M-C7	2.63927	0.74736	2.78488	0.80106	21.88178	21.65563
3E-2M-C5	2.59417	0.73012	2.77888	0.79891	20.30079	22.71023
23MM-C6	2.59179	0.72921	2.78154	0.79986	21.29856	22.38291
24MM-C6	2.63932	0.74738	2.78375	0.80066	19.93766	22.24115
33MM-C6	2.55245	0.71391	2.77026	0.79580	21.69885	22.65171
25MM-C6	2.68328	0.76390	2.78755	0.80202	19.35733	21.91039
234MMM-C5	2.61804	0.73928	2.78509	0.80114	19.13133	22.76268
233MMM-C5	2.58138	0.72518	2.78260	0.80024	19.39011	23.19700
22MM-C6	2.54201	0.70981	2.78428	0.88085	23.97131	22.25066
223MMM-C5	2.58279	0.72573	2.78380	0.80067	19.83745	23.03447
224MMM-C5	2.61183	0.73691	2.78783	0.80212	19.55118	22.60561
2233MMMM-C4	2.60260	0.73337	2.78920	0.80261	17.88167	23.56702

alkanes and are based on LOVI's, which present a "rational intramolecular ordering" of vertices.

Significant differences are observed between the ordering induced among alkanes by the various TI's presented in Tables III–VI. While VAA indices place a greater weight on branching at the penultimate carbon atoms of the longest chain (paralleling better the variation of boiling points), VAD indices place greater emphasis on central branching, paralleling better such TI's as $J^{2,9}$ or the Bertz graph derivatives.[8] The closest agreement with "ideal" intermolecular ordering is shown by the ordering induced by VAD, which is similar to that induced by the eigenvalues of $E(D)$ but differs in the following permutations for heptane and octane isomers: 2,2- vs 2,3-dimethylpentane; 2,5-dimethylhexane vs 4-methylheptane; and 2,2,4- vs 2,3,4-trimethylpentane. VAD orders the former isomer of the above three pairs as less branched, while the eigenvalues of $E(D)$ order it as more branched.

On the other hand, VEA indices in isomeric alkane series consider as the most branched alkane the linear chain with two 2,2-dimethyl side chains, which is certainly not true. Even more bizarre is the intermolecular ordering induced by VED, VRA, and VRD indices.

Some considerations about degeneracy and variation range of values will now follow. As seen above, VAA has considerable degeneracy starting with $n = 9$. VAD and VED have just one degenerate pair for $n = 9$, while VEA has no degeneracy for the alkanes with $n = 2$–9. However, VED is the least promising TI because of its senseless ordering of alkanes and its low range of variation. A low variation range is also

characteristic for VEA. By contrast, VAA, and even more so, VAD indices have a considerable range of variation, which is a useful feature for a good TI.

CORRELATIONS WITH BOILING TEMPERATURES OF ALKANES

In order to check whether the new TI's can be used in correlations with experimentally measured properties, boiling points were chosen because they are known to depend on molecular constitution (graph topology). The true nature of the intermolecular forces involved and the entropy change in the transition from liquid to gas phase are not considered in detail.

Monoparametric correlations with boiling points (at normal pressure) for all alkanes with $n = 2$–9 carbon atoms were tested for the TI's with the following results (see Chart I), where r is the correlation coefficient, s is the standard deviation, EV is the explained variance, and F is the Fisher test for 72 degrees of freedom.

It can be seen that the correlation coefficients are satisfactory for VAA1, VEA1, and VED1 and unsatisfactory for VAD1. On using the TI's labeled 3, one obtains improved r values for VAA3, VAD3, and VEA3, but a lower r value for VED3.

On using the TI's labeled 2, r values are lower than for the original TI's labeled 1 (a similar effect happens if on converting the TI's 1 into those labeled 3 one does not apply finally the logarithms). The correlation coefficients for VRA and VRD are also fairly low. An especially poor value for r is observed

REAL NUMBER VERTEX INVARIANTS

J. Chem. Inf. Comput. Sci., Vol. 31, No. 4, 1991 **521**

Table V. Data for Alkanes with 9 Carbon Atoms

alkane	BP (°C)	VAA1	VAA2	VAA3	VAD1	VAD2	VAD3
C9	150.6	5.31375	0.59042	1.56494	27.74219	3.08247	3.21759
3E-C7	143.0	5.28259	0.58695	1.55906	24.09878	2.67764	3.07680
4E-C7	141.2	5.24571	0.58286	1.55205	23.67989	2.63110	3.05927
33EE-C5	146.2	*5.23607*	0.58179	*1.55021*	20.39230	2.26581	2.90980
3M-C8	143.0	*5.23607*	0.58179	*1.55021*	25.41191	2.82355	3.12986
4E-3M-C6	140.4	5.21215	0.57913	1.54563	21.75269	2.37054	2.97438
4M-C8	142.5	5.19199	0.57689	1.54176	25.02075	2.94714	3.11434
34MM-C7	140.1	5.15642	0.57294	1.53488	22.67894	2.59188	3.01608
35MM-C7	136.0	5.14538	0.57171	1.53274	23.06872	2.56319	3.03312
3E-3M-C6	140.6	5.13056	0.57913	1.52985	21.33490	2.37054	2.95498
2M-C8	142.5	5.12583	0.56954	1.52893	26.27222	2.91914	3.16315
3E-2M-C6	138.0	5.09955	0.56662	1.52379	22.21982	2.46887	2.99562
23MM-C7	140.5	*5.08754*	*0.56528*	*1.52143*	23.55409	2.61712	3.05394
4E-2M-C6	133.8	*5.08754*	*0.56528*	*1.52143*	22.34475	2.48275	3.00123
25MM-C7	136.0	5.07524	0.56392	1.51901	23.92920	2.65880	3.06974
234MMM-C6	139.0	*5.04715*	*0.56079*	*1.51346*	21.19704	2.35523	2.94850
33MM-C7	137.3	*5.04715*	*0.56079*	*1.51346*	22.67724	2.51969	3.01600
3E-23MM-C5	141.6	5.03145	0.55905	1.51035	19.85629	2.20625	2.88316
334MMM-C6	140.5	5.01809	0.55757	1.50769	20.31723	2.25747	2.90611
24MM-C7	133.5	4.75380	0.52820	1.45358	23.54411	2.61601	3.05352
26MM-C7	135.2	4.73205	0.52578	1.44900	24.78957	2.75440	3.10506
3E-24MM-C5	136.7	*4.71557*	*0.52395*	*1.44551*	20.74384	2.30487	2.92689
44MM-C7	135.2	*4.71557*	*0.52395*	*1.44551*	22.27050	2.47450	2.99790
22MM-C7	132.7	*4.66766*	*0.51863*	*1.43530*	23.96345	2.66261	3.07117
235MMM-C6	131.3	*4.66766*	*0.51863*	*1.43530*	22.06270	2.45141	2.98853
3E-22MM-C5	133.8	*4.65028*	*0.51670*	*1.43157*	20.29814	2.25535	2.90517
233MMM-C6	137.7	*4.65028*	*0.51670*	*1.43157*	20.79452	2.31050	2.92933
223MMM-C6	131.7	4.62362	0.51374	1.42582	21.22504	2.35834	2.94982
244MMM-C6	126.5	4.59498	0.51055	1.41960	21.18387	2.35376	2.94788
2334MMMM-C5	141.5	4.57649	0.50850	1.41557	19.30054	2.14450	2.85477
224MMM-C6	126.5	4.56395	0.50711	1.41283	21.60628	2.40070	2.96762
225MMM-C6	124.0	4.52998	0.50333	1.40536	22.46618	2.49624	3.00665
2234MMMM-C5	133.0	4.51049	0.50117	1.40105	19.72565	2.19174	2.87656
2233MMMM-C5	140.3	4.49059	0.49895	1.39662	18.84402	2.09378	2.83084
2244MMMM-C5	122.7	3.96812	0.44090	1.27293	20.12632	2.23626	2.89667

Chart I

	r	s	EV	F
BP = 46.6683 VAA1 − 90.5907	0.9703	11.270	0.941	1159.3
BP = 489.471 VAA2 − 155.794	0.4427	41.789	0.185	17.6
BP = 79.7893 VAA3 + 16.1803	0.9924	5.740	0.985	4675.1
BP = 175.189 VEA1 − 342.389	0.9800	9.277	0.960	1745.3
BP = −675.334 VEA2 + 335.218	0.9600	13.051	0.920	846.1
BP = 117.386 VEA3 + 28.6100	0.9833	8.481	0.966	2102.2
BP = 7.5170 VAD1 − 28.2430	0.9407	15.814	0.833	553.3
BP = 93.9997 VAD2 − 100.8210	0.9145	18.852	0.834	368.0
BP = 56.5180 VAD3 − 36.0298	0.9814	8.941	0.963	1884.4
BP = 146.380 VED1 − 294.367	0.9895	6.743	0.979	3367.7
BP = −731.868 VED2 + 370.098	0.9603	12.998	9.921	853.7
BP = 112.353 VED3 + 24.680	0.9845	8.172	0.969	2269.6
BP = 6.426 VRA − 26.310	0.9340	16.655	0.871	491.8
BP = 6.838 VRD − 44.734	0.9639	12.413	0.928	943.0

for VAA2; this fact demonstrates that the TI's which behave well in these monoparametric correlations contain implicit information on the number n of carbon atoms and that removal of this information (conversion of TI's 1 into TI's 2) has an adverse effect on the correlation. The best correlation coefficient in these monoparametric correlations corresponds to VAA3, when the standard deviation (5.74 °C) is less than 3% of the range of BP values, and when 98.5% of the variance is accounted for by the equation. 95% confidence intervals of coefficients and Student's t 95% values for 72° of freedom are indicated by

$$BP = 79.7893(\pm 2.327) \ VAA3 + 16.1803(\pm 2.944)$$
$$(t = 68.4) \qquad\qquad (t = 11.0)$$

The next best correlations are in terms of VED1, VED3 (however, these LOVI's have a drawback to be discussed below), VEA3, and VAD3. In all these cases the standard deviation is lower than 9 °C.

Chart II

	r	s	EV	F
BP = 28.870n − 120.034	0.9850	8.034	0.970	2350.6

Chart III

	r	s	EV	F
BP = 17.007 VAA1 + 19.022n − 114.58	0.992	6.223	0.982	1984
BP = 67.353 VEA1 + 18.060n − 207.79	0.988	7.242	0.976	1456
BP = 18.593 VAD2 + 24.150n − 123.99	0.989	7.108	0.976	1512
BP = 141.68 VAA2 + 27.616n − 186.05	0.992	5.773	0.984	2311
BP = 50.375 VAA3 + 11.048n − 37.211	0.997	3.902	0.993	5101

If only the number n of carbon atoms is taken as variable, the correlation shown in Chart II results.

Biparametric linear correlations were then tested, with n and TI's as independent variables as shown in Chart III.

The last equation gives the best results, as shown by the standard deviation which is lower than 4 °C. The complete statistical parameters are shown below:

$$BP = 50.375(\pm 6.565) \ VAA3 + 11.048(\pm 2.393)n -$$
$$(t = 15.3) \qquad\qquad (t = 9.2)$$
$$37.211(\pm 11.738)$$
$$(t = 6.3)$$

One should recall that n and VAA3 are intercorrelated.

Biparametric nonlinear correlations with the same sets of two independent variables were also tested by using the following type of equation:

$$BP = b_0 + b_1 TI^x + b_2 ny; \text{ exponent values} = 0 < x, y < 1$$

For VAA1, a complete study was undertaken wherein 81 correlations (with values x and y increasing independently from 0.1 to 0.9 in steps of 0.1) were tested. We found that the highest coefficient r was obtained for $x = y = 0.5$. Therefore, we then explored all other TI's for $x = y = 0.1, 0.3, 0.5, 0.7$, and 0.9, with the results presented in Table VII. Among these

Table VI. Data for Alkanes with 9 Carbon Atoms

alkane	VEA1	VEA3	VED1	VED3	VRA1	VRD1
C9	2.82360	0.93265	*2.95052*	*0.97662*	26.66251	25.92809
3E-C7	2.71379	0.89299	2.94568	0.97498	28.01430	26.79231
4E-C7	2.71345	0.89286	2.94286	0.97402	27.33565	26.98732
33EE-C5	2.67912	0.88013	2.94338	0.97420	25.37721	28.05042
3M-C8	2.73861	0.90209	*2.95052*	*0.97662*	28.52101	26.35427
4E-3M-C6	2.71800	0.89454	2.94577	0.97501	27.00198	27.68691
4M-C8	2.72836	0.89834	2.94714	0.97548	22.84148	26.54881
34MM-C7	2.71307	0.89272	2.94731	0.97553	26.87392	27.12347
35MM-C7	2.77078	0.91377	2.95065	0.97666	25.13324	26.88862
3E-3M-C6	2.66958	0.87656	2.94391	0.97438	27.00198	27.68691
2M-C8	2.78829	0.92007	2.95372	0.97770	27.55198	26.09108
3E-2M-C6	2.71231	0.89244	2.94605	0.97511	26.51465	27.28757
23MM-C7	2.70584	0.89005	2.95066	0.97667	28.74739	26.78331
4E-2M-C6	2.76134	0.91036	2.95888	0.97945	24.95254	26.52209
25MM-C7	2.81143	0.92833	2.95416	0.97785	24.41890	26.56468
234MMM-C6	2.74126	0.90306	2.95148	0.97695	24.49290	27.43615
33MM-C7	2.65171	0.86984	2.94824	0.97585	30.30952	27.11838
3E-23MM-C5	2.69865	0.88739	2.94812	0.97581	24.42002	28.09074
334MMM-C6	2.69876	0.88743	2.94900	0.97611	25.09358	27.86608
24MM-C7	2.76300	0.91095	2.95060	0.97665	25.88491	26.78029
26MM-C7	2.85774	0.94467	2.95750	0.97898	23.65444	26.27283
3E-24MM-C5	2.73673	0.90141	2.95010	0.97648	24.19145	27.63293
44MM-C7	2.66501	0.87485	2.94483	0.97469	28.19332	27.35900
22MM-C7	2.63779	0.86458	2.95401	0.97780	34.03025	26.58714
235MMM-C6	2.77452	0.91512	2.95488	0.97810	24.06247	27.04758
3E-22MM-C5	2.70389	0.88933	2.94956	0.97629	25.13642	27.83394
233MMM-C6	2.68534	0.88245	2.94898	0.97610	26.43087	27.70301
223MMM-C6	2.68345	0.88174	2.95119	0.97685	27.72818	27.46075
244MMM-C6	2.72196	0.89599	2.95249	0.97729	24.88603	27.42730
2334MMMM-C5	2.72138	0.89578	2.95421	0.97787	23.21012	28.10802
224MMM-C6	2.74212	0.90337	2.95480	0.97807	25.25538	27.19201
225MMM-C6	2.68534	0.88245	2.94898	0.97610	25.71857	26.82626
2234MMMM-C5	2.74054	0.90280	2.95648	0.97864	23.21442	27.82602
2233MMMM-C5	2.78885	0.92027	2.96016	0.97988	23.21392	28.32479
2244MMMM-C5	2.74054	0.90280	2.95648	0.97864	22.17356	27.53977

Table VII. Correlation Coefficients for Nonlinear Equations

eq	TI	a $x = y = 0.1$	b $x = y = 0.3$	c $x = y = 0.5$	d $x = y = 0.7$	e $x = y = 0.9$
1	VAA1	0.9945	0.9970	*0.9974*	0.9960	0.9931
2	VAA2	0.9944	0.9969	*0.9974*	0.9964	0.9940
3	VAA2	0.9959	*0.9973*	0.9950	0.9915	0.9874
4	VEA1	0.9869	0.9897	*0.9905*	0.9901	0.9888
5	VEA2	0.9872	0.9904	*0.9905*	0.9901	0.9898
6	VEA3	0.9869	0.9900	*0.9905*	0.9897	0.9885
7	VRA	0.9879	0.9908	*0.9909*	0.9894	0.9869
8	VAD1	0.9882	0.9921	*0.9923*	0.9904	0.9871
9	VAD2	0.9876	0.9912	*0.9926*	0.9922	0.9900
10	VAD3	0.9897	*0.9925*	0.9913	0.9894	0.9871
11	VED1	0.9911	*0.9916*	0.9905	0.9900	0.9896
12	VED2	0.9892	0.9903	*0.9905*	*0.9905*	0.9904
13	VED3	0.9872	0.9899	0.9906	0.9910	*0.9912*
14	VRD	0.9903	0.9905	*0.9911*	0.9910	0.9902

results, the italicized ones (most of them with $x = y = 0.5$) are presented with more details in Table VIII.

The best nonlinear biparametric correlation is represented by equation 2c which is repeated below:

$$BP = -419.47(\pm19.64) + 198.64(\pm28.52) \, VAA2^{0.5} + 136.80(\pm2.59)n^{0.5}$$

From Table VI it can be seen that 99.48% of the variance of boiling points is explained by this equation, which leads to a standard deviation of only 3.35 °C.

The fact that square roots give the best results among all tested exponents is in agreement with the observation that boiling points of normal alkanes show a nonlinear increase with an increasing number of carbon atoms.

In conclusion, in this paper we have proposed as real LOVI's for graphs the eigenvectors corresponding to the lowest eigenvalues $E(A)$ and $E(D)$, and starting from them we have explored a few of the many conceivable TI's, having as main criteria the intramolecular vertex ordering for LOVI's, and

Table VIII. Coefficients of Variables in Equations from Table VII

eq	$-b_0$	b_1	b_2	r	s	F	t	EV
1c	278.19	72.77	84.93	0.9974	3.407	4405	115.0	0.9946
2c	419.47	198.64	136.80	0.9974	3.354	4546	116.8	0.9948
3b	305.73	164.32	96.77	0.9973	3.458	4275	113.2	0.9944
4c	324.09	38.30	132.71	0.9905	6.444	1214	60.4	0.9806
5c	307.59	16.21	145.40	0.9905	6.465	1206	60.2	0.9805
6c	284.72	13.66	133.62	0.9905	6.464	1206	60.2	0.9805
7c	269.75	7.29	123.51	0.9909	6.311	1267	61.7	0.9814
8c	251.76	14.37	107.28	0.9923	5.815	1496	67.0	0.9842
9c	296.72	44.85	121.10	0.9926	5.686	1567	68.6	0.9849
10b	362.92	40.94	196.61	0.9925	5.751	1531	67.8	0.9846
12c	255.94	-35.81	137.93	0.9905	6.460	1208	60.2	0.9805
12d	92.18	-158.64	58.49	0.9905	6.461	1208	60.2	0.9805
13e	211.36	-102.49	82.40	0.9912	6.208	1310	62.7	0.9820
14c	351.46	-22.72	202.32	0.9911	6.265	1286	62.1	0.9817

for TI's the intermolecular ordering of isomeric alkanes. With the new TI's we have explored for all alkanes with $n = 2-9$ carbon atoms correlations with boiling temperatures (mono-

and biparametric correlations, with the number *n* of carbon atoms in the alkane as the additional parameter; linear and nonlinear correlations, with square root as the optimal power in the latter case).

REFERENCES AND NOTES

(1) Balaban, A. T. *J. Math. Chem.*, in press.
(2) Balaban, A. T. *Pure Appl. Chem.* **1983**, *55*, 199.
(3) Lovász, L.; Pelikán, J. *Period. Math. Hung.* **1972**, *3*, 175.
(4) Bonchev, D.; Trinajstic, N. *J. Chem. Phys.* **1977**, *67*, 4517.
(5) Gutman, I.; Randic, M. *Chem. Phys. Lett.* **1977**, *47*, 15.
(6) Randič, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
(7) Randič, M.; Wilkins, C. L. *Chem. Phys. Lett.* **1979**, *63*.
(8) Bertz, S. H. *Discrete Appl. Math.* **1988**, *19*, 65.
(9) Balaban, A. T. *Chem. Phys. Lett.* **1982**, *89*, 399.
(10) Randič, M. *J. Chem. Inf. Comput. Sci.* **1975**, *15*, 105.
(11) Schultz, H. P.; Schultz, E. B.; Schultz, T. P. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 27.

Numbering of Interior Atoms in Fused Ring Systems

C. L. GLADYS* and A. L. GOODSON

Chemical Abstracts Service, P.O. Box 3012, Columbus, Ohio 43210

Because of the inadequacy of the current rules for numbering interior atoms of fused ring systems, supplementary rules have been developed at Chemical Abstracts Service. These rules are easy to apply, and in most cases, locants can be assigned by inspection.

INTRODUCTION

Existing rules for numbering interior atoms[1,2] of fused ring systems (i.e., atoms common to more than two rings) have proved inadequate. The term "clockwise numbering" is vague in the sense that it could be applied to interior atoms or to a ring system as a whole, resulting in two different sets of numbers for the same ring system. Also, no provisions were made for determining the numbering of complex branched chains of interior atoms, nor were any provisions made for numbering of interior atoms attached to a heteroatom in the interior of a fused ring system. This paper describes the supplementary rules developed at Chemical Abstracts Service (CAS) to resolve these problems, and they have been used successfully since 1980.[3]

CURRENT RULES

The current rules for assignment of orientation and numbering for fused hydrocarbons and heterocycles are applied in sequence as follows:

1. The fused ring system is oriented so that
   a. The largest number of rings are in a horizontal row.
   b. The largest number of rings are above and to the right of the horizontal row (i.e., in the upper right quadrant).
   c. The smallest number of rings are below and to the left of the horizontal row (i.e., in the lower left quadrant).
   d. The largest number of rings are above and to the left of the horizontal row (i.e., in the upper left quadrant).[4]
   e. The smallest number of rings are below and to the right of the horizontal row (i.e., in the lower right quadrant).[4]
2. If two or more orientations meet these requirements, then a unique orientation is determined as follows
   a. Low locants are given to heteroatoms.
   b. Low locants are given to heteroatoms in the priority order $O > S > Se > Te > N > P > As > Sb > Bi > Si > Ge > Sn > Pb > B$.[5]
   c. Low locants are given to carbon atoms common to two or more rings (i.e., angular carbon atoms).

d. Low locants are given to atoms that have indicated hydrogen atoms.
3. A locant is assigned to each atom of the ring system as follows
   a. Numeric locants are assigned to all non-angular, peripheral carbon atoms and to all peripheral heteroatoms, whether angular or nonangular. The locants are assigned in a clockwise direction, commencing with the most counter-clockwise, nonangular position of the uppermost ring or, if there is a choice, of the uppermost ring farthest to the right. When all of the peripheral atoms, except the angular carbon atoms, have been numbered, numbering is continued by assigning numeric locants to the interior (i.e., nonperipheral) heteroatoms.
   b. The remaining unnumbered peripheral atoms are angular carbon atoms. Such carbon atoms are designated by adding roman letters "a", "b", "c", etc., to the numeric locant of the position immediately preceding. Interior carbon atoms follow the highest numeric locant, taking a clockwise sequence when there is a choice.

SUPPLEMENTARY RULES

The supplementary rules for numbering interior angular atoms of fused ring systems replace the last sentence of Rule 3b and were developed at CAS in accordance with the principles that (1) they would be an extension of existing rules; (2) they would yield the highest allowable numbering to those atoms; (3) they would eliminate the clockwise numbering sequence of the last sentence of Rule 3b and its ambiguity. Interior atoms can occur as single atoms or as chains of atoms. A single atom is treated as if it were the first atom of a chain. A chain can be branched or unbranched. Noncontiguous chains are numbered completely and in sequence.

Where a choice is necessary because the numbering can proceed by more than one path, the following rules are applied *in sequence* until all of the interior atoms have been numbered.

Rule 1. Longest Primary Chains. Starting from the atom attached to the highest numbered angular, peripheral atom available, an unbranched chain is numbered sequentially,