ON TOPOLOGICAL INDICES INDICATING BRANCHING PART 3. ASSESSMENT OF SOME INDICES FOR THEIR SUITABILITY TO REPRESENT BRANCHING[†]

A. Perdih
Mala vas 12, SI-1000 Ljubljana, Slovenia

[†]This paper is dedicated to the late Professor Drago Kolar

Received 22-08-2000

Abstract

The susceptibility for branching, $S_{i,j}$, the difference-normalised difference of data for octanes, Δnorm_i, as well as a number of additional criteria are used to judge the suitability of 13 topological indices as branching indices. The indices EA_{max} and J are not conform with the criterion $S_{i,j}$. The other indices comply also to the criterion Δnorm_i. Due to the non-linear increase of data for n-alkanes are the most inappropriate the indices Z, W, MTI in λ_1. The most susceptible for branching are Z, D, W, MTI, Sch-S, and λ_1, whereas the least susceptible are $\lambda\lambda_1$ and ID. The same sequences of increasing branching as Tc/Pc have the indices MTI, Sch-TF, W, Xu, D, and $\lambda\lambda_1$, with $r_{Tc/Pc}$ around 0.98. No one tested index has the same sequences of increasing branching as ΔHf°g and BP/Tc. The most evenly distributed Δnorm_i data have the indices Sch-TF and D. Index D behaves similarly to W, but is not degenerated. Of BI_M and BI_A types of indices seem the most appropriate $BI_A(\chi)$ regarding $RV_{\Delta Hf°g}$, $BI_A(Xu)$ regarding $RV_{Tc/Pc}$ and $BI_A(\lambda\lambda_1)$ regarding $RV_{BP/Tc}$. BI_A indices should be included into equations of the Kamlet-Taft type.

Introduction

Several hundred topological indices have been developed and tested for their performance as branching indices or indices of substances' properties. They have been correlated with several physical, chemical, and biological properties of molecules and the interest in this has grown remarkably during the past years. Therefore, the study of branching indices remains important. Ideally, one should test all indices and correlate them with all properties. Recent studies, however, indicate that a limited number of indices may suffice for this purpose and that only few properties are suitable as references to assess the indices.

Mendiratta and Madan[1] reported that besides the Wiener[2] index (W) the most useful indices are the Hosoya[3] index (Z), the Randić[4] index (χ), and the Balaban[5] index (J). The most popular branching index, the Wiener index, W, was used even to define molecular branching[6], although it was developed to determine the paraffin boiling points[2]. Another important index is λ_1, the largest eigenvalue of the adjacency matrix[7]. At most a dozen indices emerge as the best single characterisation of diverse physicochemical properties of octanes[8].

On the basis of these findings was studied in a previous paper[9] the suitability of 11 topological indices (further on: indices) (cf. also below the chapter Data) J, W, Z, D, MTI, Xu, ID, χ, $\lambda\lambda_1$, EA_{max}, and λ_1 as branching indices, as well as of 24 physicochemical properties (further on: properties) MON, BP, d, Vi, Vm, Vc, Tc, Pc, dc, Zc, αc, ΔHv, A, B, C, n_D, MR, a_0, b_0, ΔHf°g, ΔGf°g, S, R^2, and ω, as reference properties for branching of alkanes was tested first by means of the Principal Component Analysis. It has been observed that most of properties and indices correlate highly with carbon number of alkanes although they are influenced also by branching. Of the influences of branching, assessed separately by the properties and by the indices, the most important is the number of branches, followed by the type of branched structure, i.e. whether the branch bearing carbon is tertiary or quaternary, by the position of branches, i.e. whether they are central or peripheral, distant or adjacent, and the least influence has the shape of branched molecules, i.e. whether they are spherical, flat or elongated, as well as symmetric or asymmetric. The properties have been divided into intrinsic and interaction-dependent ones and it has been explained why the latter ones are not suitable as primary references for branching. Two definitions of branching have been presented, the *Methane-based* definition as a general definition and the *n-Alkane-based* definition as a special definition more familiar to chemists[9].

Afterwards[10], additional criteria that might enable a better assessment of branching were discussed, first of all the criteria to assess the suitability of physicochemical properties as references for branching indices. The susceptibility for branching, $S_{i,j}$, the difference-normalised difference of data for octanes, $\Delta norm_i$, as well as a number of additional criteria were introduced to judge the suitability of 24 physicochemical properties of alkanes as references for branching. The most appropriate as primary references for branching regarding these criteria seemed to be ΔHf°g, Tc/Pc, and BP/Tc. They should be used simultaneously. The following properties, Tc^2/Pc, ω, and especially

C were considered to be less appropriate. Other tested physicochemical properties were found as inappropriate measures of branching. From the data of ΔHf°g, Tc/Pc, and BP/Tc the reference values for branching indices obeying the *Methane-based* definition as well as for those obeying the *n-Alkane-based* definition of molecular branching were derived[10].

In present paper we decided to study only the most frequently used indices and some recently presented topological indices (later on: indices), whereas the tests presented here can be applied also to other indices if needed. Our decision is based on the assumption that with this selection of indices no relevant information about the molecular structure contained in the information space of all indices is lost. The indices are tested in the same way as previously[10] the physicochemical properties, as well as by means of the reference values derived from ΔHf°g, Tc/Pc, and BP/Tc.

Data

The indices

We decided to take into account the group of the most frequently used indices and some novel indices. Altogether thirteen indices are used. The data for Wiener[2] index (W), the Hosoya[3] index (Z), the Randić[4] index (χ), the Balaban[5] index (J), the Yang-Xu-Hu[11] index (EA_{max}) (denoted in present paper as EA) were taken from Yang et al[11]. The ID numbers were taken from Randić[12] and the Schultz MTI index from Mihalić et al [13]. The Xu index was taken from Ren[14], whereas the Schultz indices S (Sch-S) and TLFCIR(D) (Sch-TF) were taken from Schultz and Schultz[15]. The indices $\lambda\lambda_1$[16], the largest eigenvalue of the adjacency matrix[7] (λ_1), and the largest eigenvalue of the distance matrix (D) were calculated from corresponding matrices.

The properties and the reference values for indices

They were taken from a previous paper[10].

Methods

The structures are presented in shorthand, e.g. 223M5 is 2,2,3-trimethylpentane or 3Et2M5 is 3-ethyl-2-methylpentane.

The susceptibility of indices for branching is used as the Method No. 1 to assess their suitability as branching indices. It is defined as the normalised difference of the indices' values, Eq. 1,

$$S_{i,j} = BI_i/BI_j - 1 \qquad (1)$$

where $S_{i,j}$ is the susceptibility for branching, BI is a (potential) branching index, *i* refers to the more branched structure and *j* refers to the less branched structure. Which one of the structures is more branched is concluded by intuition as presented in ref.[6] as well as by the *Methane-based* definition and *n-Alkane-based* definition[9].

Two groups of $S_{i,,j}$ data are used. In the first group, *i* refers to the structures having the same number and type of branches as *j*, but one carbon more. For example, in $S_{\mathrm{octane,heptane}}$ octane means any octane having the same number and type of branches as a heptane. In $S_{\mathrm{Oct,Hp}}$ *i* refers to *n*-octane (Oct) and *j* to *n*-heptane (Hp), or in $S_{\mathrm{2M7,2M6}}$ to 2-methyl heptane (2M7) and 2-methyl hexane (2M6), etc. In the second group of $S_{i,j}$ data, *i* and *j* refer to alkanes of the same carbon number; for example, $S_{i,\mathrm{Oct}}$ refers to any octane and *n*-octane, $S_{\mathrm{2M7,Oct}}$ refers to 2-methyl heptane and *n*-octane.

To assess the data, the following criterion is applied as the criterion No. 1: A topological index might be useful as branching index, if the sign of all $S_{\mathrm{octane,heptane}}$ values is the same and that of all $S_{i,\mathrm{Oct}}$ values is the same as well. It is not useful if the sign among members within each group differs. The two groups of sign may or may not be equal.

As the Method No. 2, the difference-normalised difference of the indices' values for octanes is used, Eq. 2,

$$\Delta\mathrm{norm}_i = (BI_i - BI_{\mathrm{Oct}})/(BI_{\mathrm{2233M4}} - BI_{\mathrm{Oct}}) \qquad (2)$$

Criterion No. 2: A topological index might be suitable as branching index if all $0 < \Delta\mathrm{norm}_i < 1$, and it is not suitable if any $\Delta\mathrm{norm}_i < 0$ or $\Delta\mathrm{norm}_i > 1$.

The criteria No. 1 and 2 serve to eliminate the indices, which are the most inappropriate candidates for good branching indices. The other criteria presented below serve to rank the remaining indices.

As the Method No. 3 we use the dependence of indices on carbon number expressed in different ways:

- The $S_{\mathrm{octane,heptane}}$ compared to the corresponding S-values derived from data for carbon number, N_C,
- The correlation coefficient with N_C, the r_{NC}.

Criterion No. 3: No clear-cut criterion could be set in this case.

As the Method No. 4 we use the estimate of the linearity of the increase of the values for *n*-alkanes with carbon number.

Criterion No. 4: The closer to linearity is its increase in *n*-alkanes with carbon number the better is an index as a branching index.

As the Method No. 5 we use the dependence of indices on branching. As a general dependence it is expressed by the $S_{i,\mathrm{Oct}}$ data as well as by the ratio $S_{i,\mathrm{Oct}}/ S_{\mathrm{octane,heptane}}$. As special indicators we use the following ratios: $S_{234M5,\mathrm{Oct}}/S_{2233M4,\mathrm{Oct}}$, $S_{25M6,\mathrm{Oct}}/S_{22M6,\mathrm{Oct}}$, and $S_{34M6,\mathrm{Oct}}/S_{33M6,\mathrm{Oct}}$ compare the susceptibility of an index for the presence of tertiary carbons vs. that of quaternary ones in octanes. The ratio $S_{25M6,\mathrm{Oct}}/S_{34M6,\mathrm{Oct}}$ indicates whether an index is more susceptible to peripheral or central position of branches.

Criterion No. 5 could not be clearly defined yet, but some useful information is given by this method.

As the Method No. 6 we use the comparison of the sequences of octanes having the same number of branches and the extent of information that can be derived thereby.

Criterion No. 6: The index having the sequences closer to those expected by intuition as well as to those of reference properties is better as a branching index.

As the Method No. 7 we use the evenness of $\Delta\mathrm{norm}_i$ data. It is assessed by the differences of $\Delta\mathrm{norm}_i$, i.e. $\Delta(\Delta\mathrm{norm}_i)_{\max}$, $\Delta(\Delta\mathrm{norm}_i)_{\min}$, as well as by their relative standard deviation (RSD). Besides the evenness as such, from these data can be deduced also the influence of the number of branches, of the (mutual) position of branches, and of the symmetry of molecules among octanes on the indices considered potentially useful as branching indices.

Criterion No. 7: The index having more evenly distributed Δnorm $_i$ data is better as a branching index.

Results and discussion

Selection of topological indices

Method No. 1: The sign of the susceptibility of indices for branching, $S_{i,,j}$
Method No. 2: The difference-normalised difference, $\Delta norm_i$

Like the physicochemical properties of alkanes[10], the indices can be assessed for their susceptibility for branching, too. Table 1 presents the grouping of indices considered here regarding the sign of $S_{i,j}$. The positive sign of $S_{i,j}$ is consistent with the fact that the indices λ_1, Sch-S, and Sch-TF increase with carbon number as well as with branching; therefore they are perspective candidates for good branching indices.

Table 1. Grouping of indices regarding the sign of their susceptibility for branching.

$S_{\text{octane,heptane}}$	$S_{i,\text{Oct}}$	Indices
All +	All +	λ_1, Sch-S, Sch-TF
All +	All -	W, D, Z, χ, ID, $\lambda\lambda_1$, MTI, Xu
All -	All +	-
All -	All -	-
Some +, some -	All +	EA, J

The indices W, D, Z, χ, ID, $\lambda\lambda_1$, MTI, and Xu increase with carbon number and decrease with branching like several properties accounted for previously[10], therefore it is not surprising that they are good indices of a single or for few those properties, but not good branching indices. On increasing carbon number, the index EA increases among *n*-alkanes, but it decreases for a series of other isomers, e.g. 2M*i*, 3M*i*, 3Et*i*, 22M*i*, 33M*i*, 23M*i*, 223M*i*, etc. ($i \geq 3$). The index J decreases in the cases of peripheral position of branches. These indices violate the *Methane-based* definition of branching and for this reason the index J[5] and especially the index EA[11] are not satisfactory as branching indices and will not be considered further. All the remaining indices comply also with the criterion No. 2, i.e. $0 < \Delta\mathrm{norm}_i < 1$.

Ranking the 11 remaining indices

Method No.3: The dependence on carbon number

To test the mean dependence of tested indices on carbon number in alkanes, the series of the mean susceptibility for the increase in carbon number from 7 to 8 at the same type of branching, $S_{\mathrm{octane,heptane}}$, of its relative standard deviation, as well as the correlation coefficient of the increase of their values with carbon number are presented in Table 2. For comparison are indicated in the series also the positions of reference properties ΔHf°g, Tc/Pc, and BP/Tc (in parentheses). According to Table 2, the indices ID, χ, Xu, $\lambda\lambda_1$, Sch-S, D, MTI, W, and especially Z are in average more susceptible for the increase in carbon number than the reference properties and N_C, whereas λ_1 is less susceptible. The relative standard deviation indicates that the spread of values of different isomers is the highest at λ_1 and the lowest at ID. The correlation with carbon number is quite high, especially at $\lambda\lambda_1$ and ID. The index Sch-TF follows the reference properties quite close.

Table 2. Susceptibility of tested indices to the increase in carbon number.

	$S_{\mathrm{octane,heptane}}$
Mean	**0.65** >Z>W~MTI>D>Sch-S>$\lambda\lambda_1$~Xu>χ~ID> (Tc/Pc) >Sch-TF~ $\boldsymbol{N_C}$ = **0.14** > (ΔHf°g)>(BP/Tc) > λ_1 > **0.01**
RSD	**0.6** > λ_1> (ΔHf°g)>(BP/Tc) > **0.1** >Z>χ> (Tc/Pc) >D~Xu~Sch-TF~MTI~W> Sch-S>$\lambda\lambda_1$ > **0.01** > ID
r_{NC}	**0.998** >$\lambda\lambda_1$>ID > **0.99** > Xu ~ Sch-TF> (Tc/Pc)>(ΔHf°g) > **0.98** >Sch-S>χ> D> MTI> **0.95** >W> (BP/Tc) > Z > λ_1> **0.85**

Method No. 4: The increase of the values for n-alkanes with carbon number

How the indices increase with carbon number in *n*-alkanes is presented in Fig. 1: Z > W > MTI >>> D > Sch-S >> $\lambda\lambda_1$ > χ > N_C > Sch-TF > Xu > ID >> λ_1. The indices Z, W, and MTI increase strongly with carbon number therefore they should be considered inferior as branching indices. Intermediate seem in this respect D and Sch-S, whereas $\lambda\lambda_1$ and χ follow N_C more closely. The indices Sch-TF and Xu increase slightly slower, whereas ID and especially λ_1 increase slower than N_C. From this point of view, the indices $\lambda\lambda_1$, χ, Sch-TF, and Xu seem to be the best of them. With exception of χ, these indices were developed recently[14-16]. It seems as if the authors (intuitively or silently) tended to develop indices approaching the linear increase with carbon number for the data of *n*-alkanes. Such a linear increase was suggested recently[10].

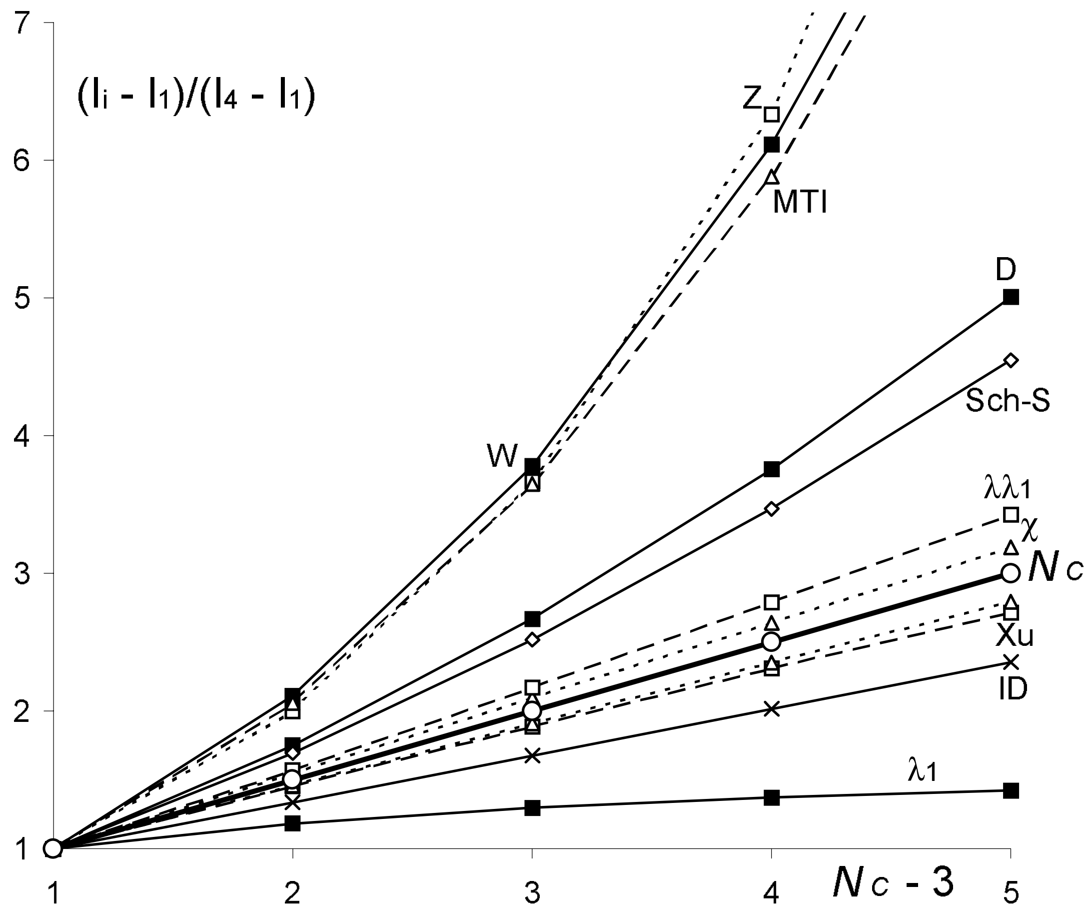

Fig. 1. The increase of indices for *n*-alkanes with carbon number (data normalised to *n*-butane). Between N_C and Xu is placed Sch-TF, marked with empty triangle.

Method No.5: The dependence on branching

The series of the mean susceptibility for the increase in branching at the same carbon number (among octanes) $S_{i,\mathrm{Oct}}$ and its relative standard deviation as well as the ratio of the mean susceptibilities is presented in Table 3. The most susceptible for branching at the same carbon number are in decreasing order Z, Sch-S, D, MTI, W, and λ_1. At this carbon number is more susceptible for branching than for the increase in carbon number only λ_1. The other indices are less susceptible for branching than for the increase in carbon number.

Table 3. Mean susceptibilities of tested indices for branching.

	$S_{i,\mathrm{Oct}}$
Mean	**0.2** > Sch-S>λ_1> **0.10** >Sch-TF> (ΔHf°g) > **0** > (BP/Tc) > **-0.02**~ID>$\lambda\lambda_1$> (Tc/Pc) >χ>Xu > **-0.1** > W~MTI~D>Z> **-0.23**
RSD	**0.65** ~ χ > (BP/Tc)> (ΔHf°g) > Z > (Tc/Pc) > ID > $\lambda\lambda_1$ > λ_1 ~ Xu > Sch-S > Sch-TF~D~W~MTI ~ **0.44**
	$S_{i,\mathrm{Oct}}$ / $S_{\mathrm{octane,heptane}}$
Mean	**10** > λ_1 >> **0.7** >Sch-TF > (ΔHf°g) > Sch-S > **0** > ID ~ $\lambda\lambda_1$ > W~MTI > Z > Xu > (Tc/Pc) > χ > **-0.4** > D > (BP/Tc) > **-1**

RSD - relative standard deviation

Instead of mean $S_{i,j}$, the $S_{i,j}$ of the most branched octane, 2233M4, as well as of other ones having a typical substitution pattern may be used for ranking the indices according to their susceptibility for branching. These rankings are presented in Table 4. The ranking according to $S_{2233\mathrm{M4,Oct}}$ indicates again that the indices Sch-S, Sch-TF, and λ_1 give rise to the same sign of $S_{i,j}$ as ΔHf°g, whereas ID, $\lambda\lambda_1$, Xu, χ, MTI, W, D, and Z give rise to the same sign of $S_{i,j}$ as Tc/Pc and BP/Tc. The ranking according to $|S_{2233\mathrm{M4,Oct}}|$ indicates that the least susceptible for branching among the indices taken in consideration are the Randić indices ID and $\lambda\lambda_1$ and that the absolute values of their susceptibilities for branching are between those of ΔHf°g and BP/Tc.

The ratios $S_{234\mathrm{M5,Oct}}/S_{2233\mathrm{M4,Oct}}$, $S_{25\mathrm{M6,8}}/S_{22\mathrm{M6,Oct}}$, and $S_{34\mathrm{M6,Oct}}/S_{33\mathrm{M6,Oct}}$ indicate that it depends on the position of branches, i.e. whether they are placed on the pepripheral or on the more centrally placed carbons, which of the indices gives more importance to tertiary carbons compared to quaternary carbons than the others. The ratio $S_{25\mathrm{M6,Oct}}/S_{34\mathrm{M6,Oct}}$

indicates that the indices Z and χ give like ΔHf°g more importance to peripheral substitution, whereas the other ones, especially λ_1, give more importance to central substitution.

Table 4. The ranking of indices regarding some susceptibilities for branching and their ratios.

Criterion	Ranking
$S_{2233M4,Oct}$	**0.24**> Sch-S>λ_1 > Sch-TF > (ΔHf°g) > **0** > (BP/Tc)>ID>$\lambda\lambda_1$ >**-0.1**> (Tc/Pc)>Xu>χ >**-0.3**> MTI>W>D > Z = **-0.5**
$\lvert S_{2233M4,Oct}\rvert$	**0.5** > Z > D>W>MTI>**0.3** > Sch-S>λ_1 > χ>Sch-TF>Xu >(Tc/Pc) > **0.1** > (ΔHf°g) > $\lambda\lambda_1$>ID>(BP/Tc)
$S_{234M5,Oct}/S_{2233M4,Oct}$	**0.8** > MTI > W>D > (Tc/Pc)>Xu > ID>$\lambda\lambda_1$>λ_1>**0.6**> Z > χ>(BP/Tc)>(ΔHf°g)
$S_{25M6,Oct}/S_{22M6,Oct}$	**0.9** > (ΔHf°g) > ID>Z>χ > Xu>D>(Tc/Pc) > MTI>W > $\lambda\lambda_1$ > **0.7** >Sch-S>(BP/Tc)> λ_1
$S_{34M6,Oct}/S_{33M6,Oct}$	(Tc/Pc)>Xu > **1.0** > MTI>D>W >Sch-TF> (BP/Tc)>$\lambda\lambda_1$>ID >Sch-S> λ_1 > **0.7** > χ > Z > (ΔHf°g)
$S_{25M6,Oct}/S_{34M6,Oct}$	**3.2** > (ΔHf°g) >Z>χ > **1** > ID>Sch-S> **0.7** > D>$\lambda\lambda_1$> Xu > Sch-TF> W > MTI>λ_1 > (BP/Tc) > (Tc/Pc) > **0.4**

Method No.6: The influence of branching of octanes on indices

To evaluate the influence of branching of octanes on indices, there are presented in Table 5 the sequences of octanes having the same number and type of branches. Several indices, i.e. MTI, Sch-TF, W, Xu, D, and $\lambda\lambda_1$ have the same sequence of octanes as Tc/Pc but the correlation of data is not what would be desired.

The correlation coefficients for data of octanes are as follows:

$r_{Tc/Pc,octanes}$: MTI (0.984) ~ Sch-TF (0.984) ~ W (0.983) ~ Xu (0.981) ~ D (0.981) >$\lambda\lambda_1$ (0.976) > λ_1 (0.965) > Sch-S (0.941) > ID (0.891) > χ (0.762) > Z (0.686)

$r_{BP/Tc,octanes}$: $\lambda\lambda_1$ (0.985) > Xu (0.972) > D (0.969) ~ W (0.968) > MTI (0.961) > ID (0.920) > χ (0.819) > Z (0.744) > Sch-S (-0.958) > Sch-TF (-0.969) > λ_1 (-0.974)

$r_{\Delta Hf°g,\ octanes}$: Sch-S (0.665) > λ_1 (0.571) > Sch-TF (0.506) > **0** > MTI (-0.467) > W (-0.506) > Xu (-0.519) > D (-0.527) > $\lambda\lambda_1$ (-0.577) > ID (-0.743) > χ (-0.849) > Z (-0.896)

Table 5. The sequences of octanes having the same number and type of branches (t - tertiary, q - quaternary). < - less branched than the next one. If not marked E for ethyl, the branch is methyl.

	Number and type of branches			
Index or property	1(t)	2(t)	2(q)	3(t) or 1(t)+2(q)
Tc/Pc and *	2<3<4<3E	25<24<23<34<3E2	22<33<3E3	224<234<223<233
BP/Tc	2<4<3<3E	25<24<23<34<3E2	22<33<3E3	224<234<223<233
$\Delta Hf°g$	3E<4<3<2	3E2<34<23<24<25	3E3<33<22	233<234<223<224
Sch-S	2<4<3<3E	25<24<23<34<3E2	22<33<3E3	234<224<233<223
λ_1	2<3<4<3E	25<24<23<34<3E2	22<33<3E3	234<224<223<233
ID	2<3E<3<4	3E2<34<25<24<23	22<33<3E3	234<224<223<233
χ	4<3E<3<2	34<3E2<23<24<25	3E3<33<22	234<233<223<224
Z	3E<3<4<2	34<3E2<23<24<25	3E3<33<22	234<233<223<224

* indices: MTI, Sch-TF, W, Xu, D, $\lambda\lambda_1$

The correlation coefficients for data of all alkanes from ethane through octanes are different, mostly appreciably higher due to the prevalent influence of carbon number:

$r_{Tc/Pc}$: Xu (0.997) > $\lambda\lambda_1$ (0.995) > ID (0.990)~ χ (0.990) > D (0.987) > MTI (0.972) > W (0.971) > Sch-TF (0.956) > Sch-S (0.949) > Z (0.915) >> λ_1 (0.777)

$r_{BP/Tc}$: Xu (0.976) > χ (0.969) ~ D (0.968) > ID (0.956) > $\lambda\lambda_1$ (0.950) > MTI (0.930) ~ W (0.929) > Z (0.886) > Sch-TF (0.881) > Sch-S (0.857) > λ_1 (0.705)

$r_{\Delta Hf°g}$: Sch-TF (0.990) > Sch-S (0.977) > ID (0.974) ~ $\lambda\lambda_1$ (0.974) > Xu (0.960) > χ (0.931) > D (0.914) > MTI (0.910) > W (0.905) > λ_1 (0.897) > Z (0.811)

The conclusion would be that the indices MTI, Sch-TF, W, Xu, D, and $\lambda\lambda_1$ represent Tc/Pc very well qualitatively but not quantitatively. The other reference properties are not represented well by these indices.

Method No. 7: The evenness of the $\Delta norm_i$ data

The evenness of the $\Delta norm_i$ data is presented here as $\Delta(\Delta norm_i)_{max}$, $\Delta(\Delta norm_i)_{min}$, and the relative standard deviation of the $\Delta norm_i$ data, RSD, Table 6. The $\Delta(\Delta norm_i)_{max}$ and RSD should be as low as possible, whereas $\Delta(\Delta norm_i)_{min}$ should be as high as possible. In this respect, the indices tend to distribute the $\Delta norm_i$ data more evenly than the best reference properties.

In the $\Delta(\Delta norm_i)_{min}$ data one weakness of some indices is indicated, namely their degeneration. Regarding the index degeneration, the data up to octanes show the degree of degeneration λ_1 > Z > W > MTI > χ, Sch-TF, whereas among octanes it is Z > W, MTI > λ_1, χ. Due to their degeneration, these indices do not seem to be appropriate ones. Having also this fact in mind, and regarding the evenness of the $\Delta norm_i$ data, then among tested indices the best ones are D and Sch-TF.

Table 6. The sequence of indices regarding the evenness of the $\Delta norm_i$ data.

Evenness	Sequence
$\Delta(\Delta norm_i)_{max}$	ID > χ > Sch-S > (BP/Tc) > $\lambda\lambda_1$ > Xu ~ D > MTI > W > λ_1 > Sch-TF > Z > (Tc/Pc) > (ΔHf°g)
$\Delta(\Delta norm_i)_{min}$	Xu > (Tc/Pc) > (BP/Tc) > D > $\lambda\lambda_1$ ~ Sch-TF > Sch-S > ID > > **0** = (ΔHf°g), χ, λ_1, MTI, W, Z
RSD	(BP/Tc) > (Tc/Pc) > (ΔHf°g) > χ > Z > ID ~ $\lambda\lambda_1$ > λ_1 ~ Xu Sch-S > Sch-TF ~ D > W > MTI

Additional remarks

The indices Z, W, and MTI increase strongly with carbon number, whereas ID and especially λ_1 increase slower. For this reason they should be considered inferior as branching indices. The indices $\lambda\lambda_1$, χ, Sch-TF, and Xu seem to be the best of them in this respect, whereas the indices D and Sch-S are intermediate. The indices χ, Sch-TF, MTI, W, and especially Z and λ_1 are degenerated. Thus, especially the indices Z, W, MTI, and λ_1 do not seem good candidates for branching indices. The index D, the largest eigenvalue of the distance matrix, behaves in many aspects similar to W but has not any degenerate value up to octanes. There is the question, why D is not used instead of W. Most indices derived from the distance matrix have a weak point, i.e. they decrease on branching. This makes them less suitable as branching indices and very suitable as interaction dependent property indices.

The index λ_1 is derived from the adjacency matrix and behaves accordingly, i.e. it increases with increasing carbon number but with a lower correlation and increases with branching. It presents some mean of the number of C-C bonds per carbon atom in alkane. It has the value of 1 for the structure containing only primary carbons (i.e. ethane). On increasing the carbon number, it asymptotically approaches the limiting value of 2 for structures containing only two primary and a number of secondary carbons. It asymptotically approaches the limiting value of 3 for structures containing also tertiary carbons, as well as the limiting value of 4 for structures containing also quaternary carbons. The simplest alkanes containing besides primary carbons also only one secondary or only one tertiary or only one quaternary carbon, have the value of λ_1 equal to the square root of the number of C-C bonds of the central carbon. Only the difference between the square root of the number of C-C bonds of the central carbon in such an alkane and the number itself is available for all the other more complex structures. Though, λ_1 indicates some interesting rules. For example, the β,ψ-dimethyl alkanes have λ_1 equal 2 regardless the size of alkane. For this reason λ_1 should not be disregarded, although as a branching index λ_1 is not what would be desired and additional indices are to be looked for, using criteria presented above.

BI_M *and* BI_A *indices derived from tested indices*

In a previous paper[10], the reference values RV to assess branching indices were derived from the properties ΔHf°g, Tc/Pc, and BP/Tc. In the same way can be derived from tested indices the candidates for branching indices that might be useful as branching indices obeying the *Methane-based* definition of branching (BI_M) or as branching indices obeying the *n-Alkane-based* definition of branching (BI_A), Eq. 3 -5.

$$BI_M = (N_C-1)^* X_i / X_n \qquad (3)$$

if X is increasing with branching, and

$$BI_M = (N_C - 1)^* X_n/X_i \qquad (4)$$

if X is decreasing with branching, where X is a topological index, *n* means the n-alkane of the same carbon number, N_C, and *i* means any alkane of the same carbon number.

The BI_A indices can be derived from BI_M indices simply[10]:

$$BI_A = BI_M - (N_C - 1) \qquad (5)$$

Another possibility would be to derive the RV_i values, as well as the BI_M and BI_A indices in the following way, Eqs. 6 - 9:

$$RV_M = (N_C\text{-}1) + \mathrm{sign}(P_i\text{-}P_n)*(P_i\text{-}P_n) \quad (6)$$

$$RV_A = \mathrm{sign}(P_i\text{-}P_n)*(P_i\text{-}P_n) \quad (7)$$

$$BI_M = (N_C\text{-}1) + \mathrm{sign}(X_i\text{-}X_n)*(X_i\text{-}X_n) \quad (8)$$

$$BI_A = \mathrm{sign}(X_i\text{-}X_n)*(X_i\text{-}X_n) \quad (9)$$

The data derived by Eq. 6 - 9 give equal or similar correlation coefficients, as those derived by Eq. 3 - 5, but their values are too dependent on the magnitude of the source data; therefore it is dissuaded from their use.

The BI_M and BI_A indices derived by Eq. 3 - 5 from the remaining 11 indices tested here were correlated with reference values derived from ΔHf°g, Tc/Pc, and BP/Tc. The results are presented in Table 7 - 9.

Table 7. Correlation coefficients of BI_M and RV_M data. Data for all alkanes from ethane through octanes are included.

BI_M from	RV_M from ΔHf°g	BI_M from	RV_M from Tc/Pc	BI_M from	RV_M from BP/Tc
χ	0.9969	Xu	0.9983	ID	0.9998
ID	0.9967	$\lambda\lambda_1$	0.9972	$\lambda\lambda_1$	0.9997
$\lambda\lambda_1$	0.9964	λ_1	0.9963	Xu	0.9955
Xu	0.9948	ID	0.9959	χ	0.9920
λ_1	0.9903	MTI	0.9945	Sch-TF	0.9898
Sch-TF	0.9889	χ	0.9937	λ_1	0.9894
W	0.9828	Sch-TF	0.9937	MTI	0.9838
D	0.9828	W	0.9930	W	0.9821
MTI	0.9824	D	0.9927	D	0.9816
Sch-S	0.9672	Sch-S	0.9686	Sch-S	0.9614
Z	0.9643	Z	0.9583	Z	0.9461

Table 8. Correlation coefficient of BI$_A$ and RV$_A$ data. Data for all alkanes from ethane through octanes are included.

BI$_A$ from	RV$_A$ from ΔHf°g	BI$_A$ from	RV$_A$ from Tc/Pc	BI$_A$ from	RV$_A$ from BP/Tc
χ	0.8920	MTI	0.9726	$\lambda\lambda_1$	0.9635
Z	0.8666	W	0.9644	D	0.9630
ID	0.8034	D	0.9633	W	0.9627
$\lambda\lambda_1$	0.7164	$\lambda\lambda_1$	0.9460	MTI	0.9626
Xu	0.7072	λ_1	0.9423	λ_1	0.9586
λ_1	0.6968	Xu	0.9278	Xu	0.9398
D	0.6537	ID	0.8945	ID	0.9216
W	0.6435	Sch-TF	0.8136	Sch-TF	0.8195
Sch-TF	0.5895	Z	0.7742	Z	0.8131
Sch-S	0.5851	χ	0.7611	χ	0.8121
MTI	0.5847	Sch-S	0.5879	Sch-S	0.6151

The correlation between BI$_M$s and RV$_M$s is high, up to over 0.999. This is not surprising since the influence of the factor having the most important contribution, N_C, was set equal. More demanding is the correlation between BI$_A$s and RV$_A$s, especially if only octanes are considered. In the case of BI$_A$s and RV$_A$s for all tested alkanes, there is highly reduced the contribution of N_C, whereas if only octanes are considered, it is entirely eliminated and the results reflect only the contribution of the influence of branching. Thus, this latter case is to be taken as the critical one in assessing the "goodness" of an index if in the former ones the index seems "good". In this respect is in the case of RV$_{\Delta Hf°g}$ the best BI$_A$ derived from χ, in the case of RV$_{Tc/Pc}$ that one derived from Xu, and in the case of RV$_{BP/Tc}$ that one derived from $\lambda\lambda_1$.

Some BI$_X$ indices (X stands for M or A) show a typical behaviour relative to the others: The correlation coefficient of BI$_X$ derived from ID decreases in all cases as the contribution of N_C decreases. The same holds true for those derived from $\lambda\lambda_1$ and λ_1 in the case of RVs derived from ΔHf°g and Tc/Pc but not BP/Tc, as well as for those derived from Xu, Sch-TF, and MTI in the case of RV$_{\Delta Hf°g}$.

The transformation of topological indices into BI$_X$ indices does not change the degeneration of those derived from χ and Sch-TF, it improves the situation with those derived from λ_1 and Z, and makes worse the situation at those derived from W and MTI. The indices Z and MTI are for this reason as well as the reasons indicated above not to be considered good indices, whereas W should be replaced by D.

Table 9. Correlation of BI_8 and RV_8 data (only data for octanes are included). BI_M and RV_M as well as of BI_A and RV_A data for octanes give rise to equal correlation coefficients.

BI_8 from	RV_8 from ΔHf°g	BI_8 from	RV_8 from Tc/Pc	BI_8 from	RV_8 from BP/Tc
Z	0.8958	Xu	0.9780	$\lambda\lambda_1$	0.9836
χ	0.8486	Sch-TF	0.9778	λ_1	0.9725
ID	0.7425	W	0.9775	Xu	0.9699
Sch-S	0.6654	MTI	0.9767	D	0.9661
$\lambda\lambda_1$	0.5773	$\lambda\lambda_1$	0.9759	Sch-TF	0.9657
λ_1	0.5708	D	0.9758	W	0.9652
D	0.5267	λ_1	0.9631	MTI	0.9576
Xu	0.5193	Sch-S	0.9393	Sch-S	0.9565
Sch-TF	0.5061	ID	0.8919	ID	0.9192
W	0.5059	χ	0.7665	χ	0.8194
MTI	0.4666	Z	0.6874	Z	0.7433

What would a good branching index be like?

Ideal index should either closely follow the increase in carbon number (in agreement with the *Methane-based* definition[9]) or should have zero value for nonbranched alkanes (*n*-alkanes; in agreement with the *n-Alkane-based* definition[9]), it should increase with branching, give for each structure a different value, etc. It should not necessarily give good single-parameter correlations with certain physicochemical properties of compounds but it should represent the branching contribution in methods like, e.g. the linear solvation energy relationship of Kamlet-Taft[17], the QSAR method of Katritzky et al.[18], and other ones. In the Kamlet-Taft correlation the size of molecules is contained in Vi whereas branching is not explicitly included. Its contribution may be hidden in different other parameters or in residual values and on inclusion of a separate parameter describing only the influence of branching the choice and the contribution of the other ones should be re-evaluated. For example, besides Vi, also BI_A should be included.

On the other hand, on inclusion of a branching index based on saturated hydrocarbons, there may be no need to consider the influence of double or triple bonds on branching index, since the influence of double or triple bonds is already accounted for in other parameters of the Kamlet-Taft equation, e.g. in π^*, α, β and analogous ones. The inclusion of an additional parameter presenting the influence of packing might be important in some cases.

References

1. S. Mendiratta, A.K. Madan, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 867-871.
2. H. Wiener, *J. Am. Chem. Soc.* **1947**, *69*, 17-20.
3. H. Hosoya, *Bull. Chem. Soc. Japan* **1971**, *44*, 2332-2339.
4. M. Randić, *J. Am. Chem. Soc.* **1975**, *97*, 6609-6615.
5. A.T. Balaban, *Chem. Phys. Lett.* **1982**, *89*, 399-404.
6. D. Bonchev, N. Trinajstić, *J. Chem. Phys.* **1977**, *67*, 4517-4533.
7. L. Lovasz, J. Pelikan, *Period. Math. Hung.* **1973**, *3*, 175-182.
8. M. Randić, S.C. Basak, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 261-266.
9. A. Perdih, M. Perdih, *Acta Chim. Slov.* **2000**, *47*, 231-259.
10. A. Perdih, *Acta Chim. Slov.* **2000**, *47*, 293-316.
11. Y.-Q. Yang, L. Xu, C.-Y. Hu, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1140-1145.
12. M. Randić, *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 164-175.
13. Z. Mihalić, S. Nikolić, N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 28-37.
14. B. Ren, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 139-143.
15. H.P. Schultz, T.P. Schultz, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 107-112.
16. M. Randić, *Acta Chim. Slov.* **1997**, *44*, 57-77.
17. M.J. Kamlet, R.M. Doherty, M.H. Abraham, Y. Marcus, R.W. Taft, *J. Phys. Chem.* **1988**, *92*, 5244-5255.
18. A.R. Katritzky, T. Tamm, Y. Wang, M. Karelson, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 692-698.

Povzetek

Občutljivost za razvejanost, $S_{i,j}$, normalizirane razlike podatkov za oktane, Δnorm_i, in še vrsta dodatnih kriterijev je bila uporabljena za vrednotenje 13 topoloških indeksov. Indeksa EA_{max} in J dajeta v isti skupini občutljivosti za razvejanost tako pozitivne kot negativne vrednosti, zato nista primerni merili razvejanosti. Drugi preizkušeni indeksi ustrezajo tej in tudi zahtevi za normalizirane razlike podatkov za oktane, Δnorm_i. Zaradi nelinearnosti naraščanja vrednosti za n-alkane so najmanj primerni indeksi Z, W, MTI in λ_1. Na razvejanje so najbolj občutljivi indeksi Z, D, W, MTI, Sch-S in λ_1, najmanj pa $\lambda\lambda_1$ in ID.

Isto zaporedje vpliva lege vej na vrednosti kot pri Tc/Pc imajo indeksi MTI, Sch-TF, W, Xu, D in $\lambda\lambda_1$, pri čemer je $r_{Tc/Pc}$ okoli 0.98. Noben preizkušen indeks nima enakega vpliva lege vej na zaporedje vrednosti kot ΔHf°g in BP/Tc. Glede na enakomernost porazdelitve vrednosti $\Delta norm_i$ sta najboljša indeksa Sch-TF in D. Indeks D naj bi uporabljali namesto indeksa W, ker nima degeneriranih vrednosti, druge lastnosti pa ima zelo podobne ali boljše kot W. Od BI_M in BI_A indeksov izvedenih iz preizkušenih indeksov so videti najprimernejši: $BI_A(\chi)$ glede na $RV_{\Delta Hf°g}$, $BI_A(Xu)$ glede na $RV_{Tc/Pc}$ in $BI_A(\lambda\lambda_1)$ glede na $RV_{BP/Tc}$. BI_A indeksi naj bi bili vključeni v enačbe kot so npr. Kamlet-Taftove.