The connectivity index is defined as
\[ \chi(G) = \sum_{u,v} (\delta_u \delta_v)^{1/2} \]
where \( \delta_v \) denotes the degree (= number of first neighbours) of the vertex \( v \) of the molecular graph \( G \), and where the summation embraces all pairs of adjacent vertices of \( G \). This structure-descriptor eventually became one of the most popular topological indices. On its applications for predicting physico-chemical and pharmacologic properties of organic compounds two books\(^2,3\) and scores of papers were written; details and further bibliography can be found in the recent monographs\(^4,6\) and the review\(^7\).

Formula (1) is a special case of a more general “connectivity index” \( C(\lambda) \), defined as:

\[ C(\lambda) = C(\lambda;G) = \sum_{u,v} (\delta_u \delta_v)^{\lambda} \]

Evidently, \( \chi = C(-0.5) \).

Several recent mathematical studies\(^8,9\) are devoted to \( C(\lambda) \), especially \( C(-1) \). We mention in passing that \( C(\lambda) \) for \( \lambda = +1 \) is the so-called “2nd Zagreb group index”; for details see the handbook\(^5\).

In the earlier paper\(^1\) no convincing argument was given why in formula (2) one should choose \( \lambda = -0.5 \).

The original aim was to provide a measure of the branching of the carbon-atom skeleton of alkanes. Based on the analysis of the data obtained for butane, pentane and hexane isomers one arrived at the conclusion\(^1\) that both choices \( \lambda = -0.5 \) and \( \lambda = -1 \) were equally plausible. Nevertheless, preference was given to the former choice because \( C(-0.5) \) had greater isomer-discriminating power than \( C(-1) \).

The possibility that the exponent \( \lambda \) assumes values other than \(-0.5\) was examined by several authors\(^10-14\). Indeed, if one would view the exponent \( \lambda \) as a variable that is adjusted so as to optimize the correlation between \( C(\lambda) \) and some physico-chemical property, then \( \lambda \) so determined would differ from \(-0.5\). The disadvantage of such an approach is that the \( \lambda \)-value so determined would significantly depend both on the physico-chemical property considered and on the sample of molecules employed.

The option \( \lambda = -1 \) was recently examined in some detail by Moon and one of the present authors\(^8\). In what follows we denote \( C(-1) = C(-1;G) \) by \( \mu = \mu(G) \).

In this paper we are concerned with certain properties of the connectivity index \( C(\lambda) \) of trees. Recall that a tree is a connected acyclic graph. A chemical tree is a tree with property \( \delta_v \leq 4 \) for all vertices \( v \). Chemical trees provide a graph representation of alkanes.

**Connectivity indices vs. branching**

Much work has been devoted to provide a precise
characterization of "branching" and to order (isomeric) molecules according to their extent of branching; see\textsuperscript{15-19} for details and further references. In spite of all these efforts there is no general agreement about what "branching" is and how it could/should be measured.

With regard to the extent of branching of trees, there is a generally accepted (and intuitively plausible) assertion: Among $n$-vertex trees the path $P_n$ and the star $S_n$ (see Fig. 1) are the least and the most branched species, respectively. In view of this, a necessary condition for any topological index to be an acceptable measure of branching is that its values be extremal for $P_n$ and $S_n$. The connectivity index satisfies this requirement: if $T$ is any $n$-vertex tree, but $T \not= P_n$, $S_n$, then:

$$\chi(P_n) > \chi(T) > \chi(S_n)$$ \hfill (3)

The proofs (which are far from being easy) of the general validity of the left- and right-hand sides of (3) were offered only a few years ago by Caporossi \textit{et al.}\textsuperscript{20} and Bollobás and Erdös\textsuperscript{21}, respectively. The path $P_n$ is, of course, a chemical tree (representing the respective normal alkane). The star $S_n$ is a chemical tree only up to $n=5$. For $n\geq 6$ also the chemical trees with minimal connectivity index were determined\textsuperscript{22}.

In the earlier paper\textsuperscript{1} it was anticipated that analogous inequalities hold when $\lambda=-1$, namely:

$$\mu(P_n) > \mu(T) > \mu(S_n)$$ \hfill (4)

As a kind of surprise, it was recently established\textsuperscript{8} that for $n$ being sufficiently large the left-hand side of (4) is violated, i.e., $P_n$ is not the tree with maximal $\mu$-value. Consequently, $\mu$ cannot be used as a measure of branching and its applicability in QSPR and QSAR studies is doubtful.

In Fig. 2 we illustrate this conclusion by comparing the correlation of $\chi = C(-0.5)$ and $\mu = C(-1)$ with the boiling points of a selected set of alkanes. In the diagram a) (that, deliberately, is identical to what earlier was reported\textsuperscript{23}), the correlation coefficient is 0.998
whereas in the diagram b) the correlation coefficient is only 0.963; more data on these correlations is found in the caption of Fig. 2.

In order to learn more on the breakdown of the inequalities (4) we determined the \( n \)-vertex tree(s) with maximal and minimal \( \mu \)-values for \( n \) up to 20. For all examined values of \( n \) the star has minimal \( \mu \)-value, in agreement with (4). For \( n \leq 9 \) the path has maximal \( \mu \)-value, in agreement with (4). However, for \( n \geq 10 \) the trees with maximal \( \mu \) differ significantly from \( P_n \); these trees are depicted in Fig. 3.

From Fig. 3, it is evident that the trees with maximal \( \mu \)-index are highly branched. Otherwise, their general structure is not easy to characterize. For \( n=16 \) and \( n=19 \) (and most probably for other values of \( n > 20 \)) the tree with maximal \( \mu \) is not unique.

Most of the species depicted in Fig. 3 are chemical trees; exceptions are only the 19-vertex tree IV and the 20-vertex tree. From the analysis outlined in the subsequent section it is evident that for \( n \) exceeding 20 more and more non-chemical trees will be encountered.

Intending to shed more light on the phenomenon described above we compared the orderings of trees according to decreasing \( \chi \) and \( \mu \), respectively. Two characteristic results are shown in Figs 4 and 5. As seen from these figures, the two orderings are significantly different, especially for larger values of \( n \) (as in Fig. 5).

The data presented in Figs 4 and 5 clearly show that the connectivity index for \( \lambda = -0.5 \) does, and the connectivity index for \( \lambda = -1 \) does not provide a plausible measure of molecular branching. In other words, when the exponent \( \lambda \) is decreased from \(-0.5\) to \(-1.0\) a drastic change occurs in the structure-dependence of the connectivity index \( C(\lambda) \), making it unsuitable for QSPR and QSAR purposes (at least, as far as branching-dependent molecular properties are concerned). In the case of \( n \)-vertex chemical trees this happens for all \( n \geq 20 \).

This change begins (or, to be more precise: becomes fully obvious) at a "critical" value of \( \lambda \) for which the equality:

\[
C(\lambda; \mathcal{T}_{crit}) = C(\lambda; P_n)
\]

is satisfied by the first \( n \)-vertex tree \( \mathcal{T}_{crit} \), different from \( P_n \). Usually (but not always) \( \mathcal{T}_{crit} \) is just one of the trees depicted in Fig. 3.

Fig. 4—The first few 10-vertex trees ordered according to decreasing connectivity indices: the species \( a_1-a_8 \) pertain to \( \chi \) whereas \( b_1-b_8 \) to \( \mu \) : \( \chi(a_1) = 4.91421, \chi(a_2) = \chi(a_3) = \chi(a_4) = 4.84606, \chi(a_5) = \chi(a_6) = \chi(a_7) = 4.80806, \chi(a_8) = 4.79475 \). \( \mu \) (b1) = 2.77778, \( \mu \) (b2) = \( \mu \) (b3) = \( \mu \) (b4) = \( \mu \) (b5) = \( \mu \) (b6) = 2.75000, \( \mu \) (b7) = \( \mu \) (b8) = 2.69444. Note that \( a_1 = b_2, a_2 = b_3, a_3 = b_4 \) and \( a_4 = b_5 \); hence in the case \( n=10 \) there is still some agreement between the two orderings.
Fig. 5—Same data as in Fig. 4, for \( n = 14 \): \( \chi(c1) = 6.91421, \chi(c2) = \chi(c3) = 6.84606, \mu(d1) = 3.81250, \mu(d2) = \mu(d3) = 3.80556, \mu(d4) = 3.80000, \mu(d5) = \mu(d6) = 3.79167 \). No tree \( c1-c9 \) coincides with any of the trees \( d1-d9 \) indicating a complete disagreement between the two orderings.

In Table 1 are given the critical values for the exponent \( \lambda \), calculated by means of Eq. (5), as well as the trees \( T_{\text{crit}} \).

From the data given in Table 1 we see that connectivity index loses its “branching index” character around \( \lambda = -0.9 \), which is relatively far from the adopted value \( \lambda = -0.5 \) and relatively near to \( \lambda = -1 \). The original choice of \( \lambda = -0.5 \) for the exponent in the definition of the connectivity index (instead of \( \lambda = -1 \)) is now seen to be a rather fortunate decision. This certainly is one of the reasons of the great quarter-of-century-long success of the connectivity index. Our analysis sheds some new light on the true meaning of this choice.

**Beyond \( \lambda = -1 \)**

The cause of the above described “breakdown” of the connectivity index is best seen if we consider the behaviour of \( C(\lambda) \) for large negative values of \( \lambda \). For this, denote by \( m_{ij} \) the number of edges connecting a vertex of degree \( i \) with a vertex of degree \( j \) and rewrite Eq. (2) as:

\[
C(\lambda) = \sum_{i \leq j} m_{ij}(i:j)^{\lambda} \quad \ldots \quad (6)
\]

Recall that the condition \( m_{11} = 0 \) is satisfied by all connected graphs except by the molecular graph of ethane. Therefore, in what follows we assume that the term \( m_{11} (1:1)^{\lambda} = m_{11} \) is not contained among the summands on the right-hand side of (6).

Now, for large negative values of the exponent \( \lambda \) all summands on the right-hand side of (6) become small and will vanish as \( \lambda \rightarrow -\infty \). The term that decreases most slowly will thus become dominant. This term, evidently, is \( m_{12} 2^{\lambda} \), implying that for large negative values of \( \lambda \):

\[
C(\lambda) = m_{12} 2^{\lambda}
\]

This means that for large negative values of \( \lambda \), \( C(\lambda) \) will be maximal for the graph possessing the greatest...
possible number of edges that connect a vertex of degree one with a vertex of degree two. The trees with such a property are \( X_n \) depicted in Fig. 1.

By means of Eq. (6) it is easy to show that the tree with minimal \( C(\lambda) \)-value, no matter how close \( \lambda \) is to \(-\infty\), is the star \( S_n \). Thus, for large negative values of \( \lambda \) instead of the relation (3) we have:

\[
C(\lambda; X_n) > C(\lambda; T) > C(\lambda; S_n)
\]

As seen in Fig. 1, both trees \( X_n \) and \( S_n \) are highly branched. Consequently, when \( \lambda \) is sufficiently large (and negative) then the ordering of trees (and chemical trees) by means of the connectivity index \( C(\lambda) \) is not determined by the extent of branching, but by some quite different (above specified) structural features. Consequently, with \( \lambda \) becoming more and more negative its “branching index” character must necessarily disappear. As shown in the preceding section, for \( n \geq 10 \) this happens before \( \lambda \) becomes equal to \(-1\).

The alkanes (chemical trees) that for large negative \( \lambda \) have the greatest connectivity index are also easily characterized. These are specified in Table 2. The main structural detail determining these extremal systems is the presence of as many as possible ethyl groups (edges connecting vertices of degree one and two). The pattern by which they are constructed is readily envisaged: it has a periodicity of length five so that for all \( n=5, 6, \ldots \), the construction step \( n \to n+1 \) is fully analogous to the step \( n+5 \to n+6 \).

References
1 Randić M, J Am chem Soc, 97 (1975) 6609.