INFORMATION MEASURES TO CHARACTERIZE WEIGHTED CHEMICAL STRUCTURES

Matthias Dehmer and Stephan Borgert

Abstract. To characterize molecule structures by using information-theoretic techniques is an interesting and challenging problem in mathematical chemistry. However, most of the classical information indices are only defined for characterizing the skeletons of chemical graphs that correspond to unweighted graphs. This work presents a possible extension of an information measure that was recently developed to determine the structural information content of unweighted graphs. The novel measure then takes edge and vertex labels into account when measuring the information content of a weighted network. We illustrate the calculation of the resulting information measure by considering a special weighted chemical graph.

Keywords: *Information Measures, Complex Systems, Structural Complexity, Labeled Chemical Graphs, Information Theory, Chemical Graph Theory*

1 Introduction

A major part of the research in chemical graph theory deals with investigating structural properties of molecules representing graphs by using so-called topological and information-theoretic indices [3, 2, 12, 13]. The latter is the object of research we cover in the present paper. Generally, information theory [15] offers quantitative methods to investigate, e.g., information processing and information transmission in graphs [2]. Particularly in mathematical chemistry, many problems in QSAR and QSPR require methods for analyzing structural properties of chemical graphs quantitatively. QSAR (Quantitative structure-activity relationship) deals with descripting pharmacokinetic processes as well

as biological activity or chemical reactivity [1, 8]. In contrast, QSPR (Quantitative Structure-Property Relationship) addresses the problem to convert chemical structures into molecular descriptors [17] which are relevant to a physico-chemical property or a biological activity [8, 9]. Especially, a main problem in QSPR is to investigate relationships between molecular structure and physicochemical properties [2].

In [6], we recently introduced a novel information measure to determine the structural information content of unlabeled and undirected chemical graphs (see Section (3)). So far, this entropic measure has been mainly used to detect molecular branching in molecules representing the just mentioned graph class [7].

As a main contribution of our paper, we want to extend the information measure presented in [7] to vertex- and edge-labeled (weighted) graphs because chemical structures can be adequately represented by graphs only if different types of atoms (vertices) and different types of bonds (edges) are considered. For instance, heteroatoms like nitrogen (N) or oxygen (O) instead of carbon (C) change many properties of the molecules considerably. Double bonds and triple bonds are much more reactive than single bonds.

A classical contribution to quantify the amount of information for the kind of atoms in a molecule was given by [2, 5]. By using the total or mean information [2], it turned out that the so-called information indices on atomic composition for molecules denoted by the empirical formulas $A_xB_yC_z$ can be determined as [2, 5]

$$I^t_{AC} := (x+y+z)\log(x+y+z) - x\log(x) - y\log(y) - z\log(z), \quad (1)$$

or

$$I^m_{AC} := -p_x\log(p_x) - p_y\log(p_y) - p_z\log(p_z), \quad (2)$$

where

$$p_x := \frac{x}{x+y+z},\ p_y := \frac{z}{x+y+z}, \quad \text{and} \quad p_z := \frac{z}{x+y+z}. \quad (3)$$

However, we see that these indices do not take into account any structural properties of the molecules. In contrast, in the following we put the emphasis on developing an information measure for characterizing chemical structures representing vertex- and edge-labeled graphs.

2 Mathematical Preliminaries

In this section, we briefly introduce mathematical preliminaries [4, 11, 10] to formulate our approach. For this, we first start with the definition of finite, undirected and connected graphs. We call $G = (V, E), |V| < \infty$, $E \subseteq \binom{V}{2}$ a finite, undirected and connected graph whereas $\mathcal{G}_{UC}$ denotes the set of such graphs. In the following, we also repeat the definitions of some metrical properties of graphs [7, 16]. $d(u, v)$ denotes the shortest distance between $u \in V$ and $v \in V$. $d(u, v)$ is an integer metric. For $G \in \mathcal{G}_{UC}$, $\sigma(v) = \max_{u \in V} d(u, v)$ is called the eccentricity of $v \in V$ and $\rho(G) = \max_{v \in V} \sigma(v)$ the diameter of G, respectively.

$$S_j(v_i, G) := \{v \in V | \, d(v_i, v) = j, \, j \geq 1\}, \tag{4}$$

denotes the j-sphere of v_i regarding $G \in \mathcal{G}_{UC}$. Further, we state the following definitions.

Definition 2.1 *Let*

$$A_V^G := \{l_v^1, l_v^2, \ldots, l_v^{|A_V^G|}\}, \tag{5}$$

and

$$A_E^G := \{l_e^1, l_e^2, \ldots, l_e^{|A_E^G|}\}, \tag{6}$$

be unique (finite) vertex and edge alphabets, respectively. $l_V : V \longrightarrow A_V^G$ *and* $l_E : E \longrightarrow A_E^G$ *are the corresponding edge and vertex labeling functions. Then, we call* $G := (V, E, l_V, l_E)$ *a finite, undirected and labeled graph.* $\mathcal{G}_{UL}$ *denotes the set of finite, undirected and labeled graphs.*

Definition 2.2 *Let* X *be a discrete random variable with alphabet* A *and* $p(x_i) = Pr(X = x_i)$ *be the probability mass function of* X*. Then, the entropy of* X *is defined by*

$$H(X) := -\sum_{x_i \in A} p(x_i) \log(p(x_i)). \tag{7}$$

3 Structural Information Content of Unweighted Graphs

We briefly repeat the construction of the entropy measure for determining the structural information content of unlabeled graphs that was recently introduced in [7, 6]. In [7], the resulting information measure was mainly used to

detect molecular branching in unlabeled chemical graphs. As mentioned, the chemical graphs were considered as skeletons only, i.e., all atoms and all bonds were considered as equal. To define a probability distribution by inferring structural characteristics from the graphs under consideration, we used the definition of the j-spheres, see Equation (4). Finally, we defined the quantities

$$p^V(v_i) := \frac{f^V(v_i)}{\sum_{j=1}^{|V|} f^V(v_j)}, \tag{8}$$

which have been interpreted as vertex probabilities [7]. Further,

$$f^V(v_i) := \alpha^{c_1|S_1(v_i,G)|+c_2|S_2(v_i,G)|+\cdots+c_{\rho(G)}|S_{\rho(G)}(v_i,G)|}, \\ c_k > 0, 1 \leq k \leq \rho(G), \alpha > 0, \tag{9}$$

represents an information functional that captures structural information of a graph. c_k are arbitrary real positive coefficients to be chosen such that they are not all equal. We see that this information functional is based on the inferred j-sphere cardinalities. From these definitions, the structural information content of $G \in \mathcal{G}_{UC}$ has been defined as its corresponding entropy [6]. As a result, we obtained

$$I_{f^V}(G) := -\sum_{i=1}^{|V|} \frac{f^V(v_i)}{\sum_{j=1}^{|V|} f^V(v_j)} \log \left(\frac{f^V(v_i)}{\sum_{j=1}^{|V|} f^V(v_j)} \right), \tag{10}$$

that represents a family of graph entropy measures. From a mathematical viewpoint, it turned out [7] that this entropy measure generalizes most of the classical information indices used in mathematical and computational chemistry [2, 14, 17].

4 Structural Information Content of Weighted Graphs

In this section, we want to extend the proposed entropy measure from the previous section for determining the structural information content of vertex- and edge-labeled (weighted) graphs. Let $G = (V, E) \in \mathcal{G}_{UL}$ and let $v_i \in V$ be an arbitrary vertex. Further, we assume that G is arbitrarily labeled (regarding

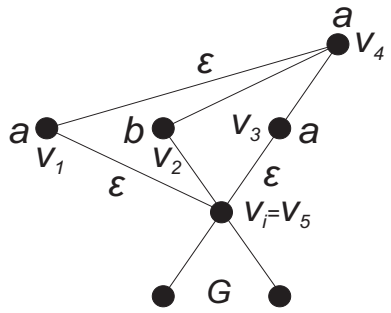

Figure 1: A special vertex- and edge-labeled graph.

the labels $v_i, 1 \leq i \leq |V|$). In the following, we define the paths

$$P_1^j := (v_i, v_{a_1}, v_{a_2}, \ldots, v_{a_j}), \tag{11}$$
$$P_2^j := (v_i, v_{b_1}, v_{b_2}, \ldots, v_{b_j}), \tag{12}$$
$$\vdots$$
$$P_{k_j}^j := (v_i, v_{x_1}, v_{x_2}, \ldots, v_{x_j}), \tag{13}$$

which are induced by determining the shortest paths starting from v_i. Hence, the corresponding edge sets are given by

$$E_1^j := \{\{v_i, v_{a_1}\}, \{v_{a_1}, v_{a_2}\}, \ldots, \{v_{a_{j-1}}, v_{a_j}\}\}, \tag{14}$$
$$E_2^j := \{\{v_i, v_{b_1}\}, \{v_{b_1}, v_{b_2}\}, \ldots, \{v_{b_{j-1}}, v_{b_j}\}\}, \tag{15}$$
$$\vdots$$
$$E_{k_j}^j := \{\{v_i, v_{x_1}\}, \{v_{x_1}, v_{x_2}\}, \ldots, \{v_{x_{j-1}}, v_{x_j}\}\}, \tag{16}$$

and we set

$$E^j := E_1^j \cup E_2^j \cup \cdots \cup E_{k_j}^j. \tag{17}$$

As an example, we consider $G = (V, E) \in \mathcal{G}_{UL}$ depicted in Figure (1). It holds

$$l_v(v_1) = l_v(v_3) = l_v(v_4) = a,\ l_v(v_2) = b, \tag{18}$$

and

$$l_e(\{v_i, v_1\}) = l_e(\{v_i, v_3\}) = l_e(\{v_1, v_4\}) = \epsilon, \tag{19}$$

where $A_V^G = \{a, b\}$ and $A_E^G = \{\epsilon\}$. Further, to determine the paths induced by the 2-sphere exemplarily, we obtain

$$P_1^2 := (v_i, v_{a_1}, v_{a_2}), \tag{20}$$
$$P_2^2 := (v_i, v_{b_1}, v_{b_2}), \tag{21}$$
$$P_3^2 := (v_i, v_{c_1}, v_{c_2}), \tag{22}$$

and

$$v_{a_1} = v_3,\ v_{b_1} = v_2,\ v_{c_1} = v_1,\ v_{a_2} = v_{b_2} = v_{c_2} = v_4. \tag{23}$$

We want to emphasize that if the vertex numbering labels v_i change, the vertex and edge labels remain unchanged. Therefore, for determining the paths as outlined, it suffices to consider the corresponding unlabeled version of this graph. Now, for $j = 1, 2, \ldots, \rho(G)$, we define the quantities

$$|S_j^{l_v^\mu}(v_i, G)| := |\{v \in V |\, d(v_i, v) = j,\ j \geq 1,\ l_V(v) = l_v^\mu, \\ \mu = 1, 2, \ldots, |A_V^G|\}|, \tag{24}$$

and

$$|E_{l_e^\mu}^j| := |\{e \in E_1^j \cup E_2^j \cup \cdots \cup E_{k_j}^j |\, e \text{ is incident with } v \in S_j(v_i, G) \\ \wedge\ l_E(e) = l_e^\mu,\ \mu = 1, 2, \ldots, |A_E^G|\}|. \tag{25}$$

To illustrate the given definitions, we consider Figure (2). This figure shows a special chemical structure represented by a vertex- and edge-labeled chemical graph G. We set $A_V^G = \{\mathrm{O}, \mathrm{C}, \mathrm{N}\}$ and $A_E^G = \{s, d\}$. O, C and N denote the kind of atoms. The edge type s represents a single bond whereas d represents a double bond. For example, if we choose v_3 as a starting vertex, we yield

$$|S_1^O(v_3, G)| = 1, |S_1^C(v_3, G)| = 1, |S_1^N(v_3, G)| = 1, \tag{26}$$
$$|S_2^O(v_3, G)| = 0, |S_2^C(v_3, G)| = 2, |S_2^N(v_3, G)| = 0, \tag{27}$$
$$|S_2^O(v_3, G)| = 1, |S_2^C(v_3, G)| = 0, |S_2^N(v_3, G)| = 0, \tag{28}$$

and

$$|E_s^1| = 2, |E_d^1| = 1, \tag{29}$$
$$|E_s^2| = 2, |E_d^2| = 0, \tag{30}$$
$$|E_s^3| = 1, |E_d^3| = 0. \tag{31}$$

By using the above given definitions, we first obtain direct generalizations of the information functional expressed by Equation (9).

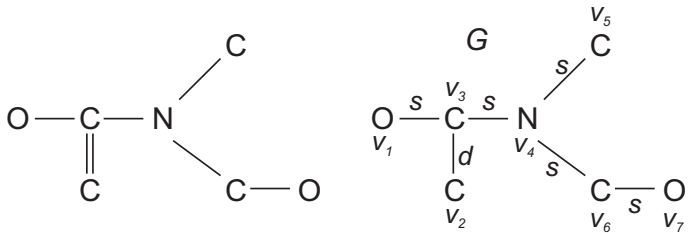

Figure 2: Left: A chemical structure representing $C_4H_9NO_2$. Right: The corresponding vertex- and edge-labeled graph G.

Definition 4.1 *Let $G = (V, E) \in \mathcal{G}_{UL}$ and we assume that $A_V^G \neq \emptyset$. We define*

$$f_l^V(v_i) := \alpha^{\sum_{k=1}^{\rho(G)} \sum_{\mu=1}^{|A_V^G|} c_k^{l_v^\mu} |S_k^{l_v^\mu}(v_i,G)|}, \; c_k^{l_v^\mu} > 0, \; \alpha > 0. \tag{32}$$

Definition 4.2 *Let $G = (V, E) \in \mathcal{G}_{UL}$ and we assume that $A_E^G \neq \emptyset$. We define*

$$f_l^E(v_i) := \alpha^{\sum_{k=1}^{\rho(G)} \sum_{\mu=1}^{|A_E^G|} b_k^{l_e^\mu} |E^j_{l_e^\mu}|}, \; b_k^{l_e^\mu} > 0, \; \alpha > 0. \tag{33}$$

Now, it is straightforward to derive an information functional for quantifying structural information of vertex- and edge-labeled graphs.

Definition 4.3 *Let $G = (V, E) \in \mathcal{G}_{UL}$ and we assume that $A_V^G, A_E^G \neq \emptyset$. We define*

$$f_l^{V,E}(v_i) := \alpha^{\sum_{k=1}^{\rho(G)} \sum_{\mu=1}^{|A_V^G|} c_k^{l_v^\mu} |S_k^{l_v^\mu}(v_i,G)| + \sum_{k=1}^{\rho(G)} \sum_{\mu=1}^{|A_E^G|} b_k^{l_e^\mu} |E^j_{l_e^\mu}|}, \\ b_k^{l_e^\mu}, c_k^{l_v^\mu} > 0, \; \alpha > 0. \tag{34}$$

From this, we yield the corresponding families of entropic measures to determine the structural information content of labeled graphs.

Definition 4.4 *Let $G = (V, E) \in \mathcal{G}_{UL}$ and we assume that $A_V^G, A_E^G \neq \emptyset$. We obtain the following information measures:*

$$I_{f_l^V}(G) := -\sum_{i=1}^{|V|} \frac{f_l^V(v_i)}{\sum_{j=1}^{|V|} f_l^V(v_j)} \log\left(\frac{f_l^V(v_i)}{\sum_{j=1}^{|V|} f_l^V(v_j)}\right), \tag{35}$$

$$I_{f_l^E}(G) := -\sum_{i=1}^{|V|} \frac{f_l^E(v_i)}{\sum_{j=1}^{|V|} f_l^E(v_j)} \log\left(\frac{f_l^E(v_i)}{\sum_{j=1}^{|V|} f_l^E(v_j)}\right), \tag{36}$$

$$I_{f_l^{V,E}}(G) := -\sum_{i=1}^{|V|} \frac{f_l^{V,E}(v_i)}{\sum_{j=1}^{|V|} f_l^{V,E}(v_j)} \log\left(\frac{f_l^{V,E}(v_i)}{\sum_{j=1}^{|V|} f_l^{V,E}(v_j)}\right). \tag{37}$$

To determine the information functionals $f_l^V, f_l^{V,E}$ as well as the corresponding structural information contents exemplarily, we again consider Figure (2). First, we determine the information functional $f_l^{V,E}(v_i)$ as follows:

$$f^{V,E}(v_1) := \alpha^{c_1^{\mathrm{C}}+c_2^{\mathrm{C}}+c_2^{\mathrm{N}}+2c_3^{\mathrm{C}}+c_4^{\mathrm{O}}+b_1^s+b_2^s+b_2^d+2b_3^s+b_4^s}, \tag{38}$$

$$f^{V,E}(v_2) := \alpha^{c_1^{\mathrm{C}}+c_2^{\mathrm{O}}+c_2^{\mathrm{N}}+2c_3^{\mathrm{C}}+c_4^{\mathrm{O}}+b_1^d+2b_2^s+2b_3^s+b_4^s}, \tag{39}$$

$$f^{V,E}(v_3) := \alpha^{c_1^{\mathrm{O}}+c_1^{\mathrm{C}}+c_1^{\mathrm{N}}+2c_2^{\mathrm{C}}+c_3^{\mathrm{O}}+2b_1^s+b_1^d+2b_2^s+b_3^s}, \tag{40}$$

$$f^{V,E}(v_4) := \alpha^{3c_1^{\mathrm{C}}+2c_2^{\mathrm{O}}+c_2^{\mathrm{C}}+3b_1^s+2b_2^s+b_3^d}, \tag{41}$$

$$f^{V,E}(v_5) := \alpha^{c_1^{\mathrm{N}}+2c_2^{\mathrm{C}}+2c_3^{\mathrm{O}}+c_3^{\mathrm{C}}+b_1^s+2b_2^s+b_3^d+2b_3^s}, \tag{42}$$

$$f^{V,E}(v_6) := \alpha^{c_1^{\mathrm{O}}+c_1^{\mathrm{N}}+2c_2^{\mathrm{C}}+c_3^{\mathrm{O}}+c_3^{\mathrm{C}}+2b_1^s+2b_2^s+b_3^d+b_3^s}, \tag{43}$$

$$f^{V,E}(v_7) := \alpha^{c_1^{\mathrm{N}}+c_2^{\mathrm{N}}+2c_3^{\mathrm{C}}+c_4^{\mathrm{O}}+c_4^{\mathrm{C}}+b_1^s+b_2^s+2b_3^s+b_4^d+b_4^s}. \tag{44}$$

To obtain these equations, we first compute the 1-sphere cardinalities for vertices having the labels O, C and N

$$|S_1^C(v_i, G)|, |S_1^O(v_i, G)|, |S_1^N(v_i, G)|, \tag{45}$$

and multiply those with the corresponding coefficients $c_1^{\mathrm{C}}, c_1^{\mathrm{O}}, c_1^{\mathrm{N}}$. Then, we perform the same step to calculate

$$S_2^C(v_i, G)|, |S_2^O(v_i, G)|, |S_2^N(v_i, G)|. \tag{46}$$

For each vertex v_i, we continue this step until $j = \rho(G) = 4$ holds. Second, for each vertex v_i, we compute the quantities $|E_s^1|$ and $|E_d^1|$, i.e., the number

of edges induced by $P_1^1, P_2^1, \ldots, P_{k_1}^1$ having the edge types s and d, respectively. After this calculation, we multiply the computed quantities with the corresponding coefficients b_1^s and b_1^d. We continue this step until calculating $E_s^{\rho(G)}|$ and $|E_d^{\rho(G)}|$ induced by $P_1^{\rho(G)}, P_2^{\rho(G)}, \ldots, P_{k_{\rho(G)}}^{\rho(G)}$. As a result, we obtained Equation (38)-(44).
For example, if we now set

$$c_1^{\mathrm{O}} := 9 > c_2^{\mathrm{O}} := 7 > c_3^{\mathrm{O}} := 5 > c_4^{\mathrm{O}} := 2, \tag{47}$$
$$c_1^{\mathrm{C}} := 8 > c_2^{\mathrm{C}} := 6 > c_3^{\mathrm{C}} := 4 > c_4^{\mathrm{C}} := 2, \tag{48}$$
$$c_1^{\mathrm{N}} := 6 > c_2^{\mathrm{N}} := 4 > c_3^{\mathrm{N}} := 2 > c_4^{\mathrm{N}} := 1, \tag{49}$$
$$b_1^{\mathrm{s}} := 8 > b_2^{\mathrm{s}} := 6 > b_3^{\mathrm{s}} := 4 > b_4^{\mathrm{s}} := 2, \tag{50}$$
$$b_1^{\mathrm{d}} := 6 > b_2^{\mathrm{d}} := 4 > b_3^{\mathrm{d}} := 2 > b_4^{\mathrm{d}} := 1, \tag{51}$$

we yield

$$f^{V,E}(v_1) := \alpha^{57}, \tag{52}$$
$$f^{V,E}(v_2) := \alpha^{58}, \tag{53}$$
$$f^{V,E}(v_3) := \alpha^{78}, \tag{54}$$
$$f^{V,E}(v_4) := \alpha^{82}, \tag{55}$$
$$f^{V,E}(v_5) := \alpha^{56}, \tag{56}$$
$$f^{V,E}(v_6) := \alpha^{70}, \tag{57}$$
$$f^{V,E}(v_7) := \alpha^{50}. \tag{58}$$

Finally, the structural information content of G by using the computed information functionals $f_l^{V,E}(v_i)$ becomes to

$$\begin{aligned} I_{f_l^{V,E}}(G) := -\sum_{i=1}^{|V|} \frac{\alpha^{50}}{D} \log\left(\frac{\alpha^{50}}{D}\right) + \frac{\alpha^{56}}{D} \log\left(\frac{\alpha^{56}}{D}\right) + \frac{\alpha^{57}}{D} \log\left(\frac{\alpha^{57}}{D}\right) \\ + \frac{\alpha^{58}}{D} \log\left(\frac{\alpha^{58}}{D}\right) + \frac{\alpha^{70}}{D} \log\left(\frac{\alpha^{70}}{D}\right) + \frac{\alpha^{78}}{D} \log\left(\frac{\alpha^{78}}{D}\right) \\ + \frac{\alpha^{82}}{D} \log\left(\frac{\alpha^{82}}{D}\right), \end{aligned} \tag{59}$$

where

$$D := \alpha^{50} + \alpha^{56} + \alpha^{57} + \alpha^{58} + \alpha^{70} + \alpha^{78} + \alpha^{82}. \tag{60}$$

Now, Equation (59) represents a family of information measures for measuring the structural information content of the vertex- and edge-labeled chemical graph G shown in Figure (2). Further, we see that Equation (59) now depends on a free parameter α only. As a result, we obtain different entropy values by varying this parameter. Finally, this gives us a possibility to study the local information spread in chemical graphs.

5 Summary and Conclusion

In this paper, we presented a generalization of a recently developed method to determine the structural information content of unlabeled chemical graphs [7]. The structural information content has been interpreted as the entropy of the underlying graph topology. Our generalization done in this paper concerns the problem to measure the entropy of vertex- and edge-labeled chemical graphs. As a result, we got certain families of entropy measures to characterize labeled chemical graphs. As a result, we see that the proposed approach based on using certain information functionals is powerful because different information functionals can be easily incorporated. That means, to determine the structural information content of special labeled chemical graphs, special information functionals can be defined. Further, starting from appropriate data sets, the idea of incorporating parameterized information functionals offers the perspective to use machine learning methods for learning optimal parameters. In particular, this can be interesting for predicting the entropy of labeled chemical graphs by using chemical databases.

As future work, we want to apply our method to large databases for classifying labeled chemical structures automatically.

6 Acknowledgments

We would like to thank Danail Bonchev, Frank Emmert-Streib, Abbe Mowshowitz and Kurt Varmuza for fruitful discussions.

References

[1] R. Benigni. *Quantitative Structure-Activity Relationship (QSAR) Models of Mutagens and Carcinogens*. CRC Press, 2003.

[2] D. Bonchev. *Information Theoretic Indices for Characterization of Chemical Structures*. Research Studies Press, Chichester, 1983.

[3] D. Bonchev and N. Trinajstić. Information theory, distance matrix and molecular branching. *Journal of Chemical Physics*, 67:4517–4533, 1977.

[4] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications and Signal Processing. Wiley & Sons, 2006.

[5] S. M. Dancoff and H. Quastler. Information content and error rate of living things. In H. Quastler, editor, *Essays on the Use of Information Theory in Biology*, pages 263–274. University of Illinois Press, 1953.

[6] M. Dehmer. A novel method for measuring the structural information content of networks. *Cybernetics and Systems*, 39:825–843, 2008.

[7] M. Dehmer and F. Emmert-Streib. Structural information content of chemical networks. *Zeitschrift für Naturforschung, Part A*, 63a:155–159, 2008.

[8] J. Devillers and A. T. Balaban. *Topological Indices and Related Descriptors in QSAR and QSPR*. Gordon and Breach Science Publishers, 1999. Amsterdam, The Netherlands.

[9] M. V. Diudea. *QSPR / QSAR Studies by Molecular Descriptors*. Nova Publishing, 2001.

[10] R. Halin. *Graphentheorie*. Akademie Verlag, 1989. Berlin, Germany.

[11] F. Harary. *Graph Theory*. Addison Wesley Publishing Company, 1969. Reading, MA, USA.

[12] E. V. Konstantinova. On some applications of information indices in chemical graph theory. In R. Ahlswede, L. Bäumer, N. Cai, H. Aydinian, V. Blinovsky, C. Deppe, and H. Mashurian, editors, *General Theory of Information Transfer and Combinatorics*, Lecture Notes of Computer Science, pages 831–852. Springer, 2006.

[13] M. Randić and D. Plavšić. On the concept of molecular complexity. *Croatica Chemica Acta*, 75:107–116, 2002.

[14] B. M. Rode, T. S. Hofer, and M. D. Kugler. *The Basics of Theoretical and Computational Chemistry*. Wiley-VCH, 2007.

[15] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, 1997. Urbana, IL, USA.

[16] V. A. Skorobogatov and A. A. Dobrynin. Metrical analysis of graphs. *Commun. Math. Comp. Chem.*, 23:105–155, 1988.

[17] R. Todeschini, V. Consonni, and R. Mannhold. *Handbook of Molecular Descriptors*. Wiley-VCH, 2002. Weinheim, Germany.

Matthias Dehmer, Discrete Mathematics and Geometry, Vienna University of Technology, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria,
email: *mdehmer@dmg.tuwien.ac.at*

Stephan Borgert, Darmstadt University of Technology, Hochschulstr. 10, 64289 Darmstadt, Germany,
email: *borgert@tk.informatik.tu-darmstadt.de*