Chapter 2

Graph–Theoretical Indices based on Simple, General and Complete Graphs

Lionello Pogliani
Università della Calabria, Italy

ABSTRACT

Valence molecular connectivity indices are based on the concept of valence delta, δ^v*, that can be derived from general chemical graphs or chemical pseudographs. A general graph or pseudograph has multiple edges and loops and can be used to encode, through the valence delta, chemical entities. Two graph-theoretical concepts derived from chemical pseudographs are the intrinsic (I) and the electrotopological state (E) values, which are the used to define the valence delta of the pseudoconnectivity indices,* $\psi_{I,S}$*. Complete graphs encode, through a new valence delta, the core electrons of any atoms in a molecule. The connectivity indices, either valence connectivity or pseudoconnectivity, are the starting point to develop the dual connectivity indices. The dual indices show that not only can they assume negative values but also cover a wide range of numerical values. The central parameter of the molecular connectivity theory, the valence delta, defines a completely new set of connectivity indices, which can be distinguished by their configuration and advantageously used to model different properties and activities of compounds.*

INTRODUCTION

The many new types of molecular connectivity indices (Pogliani, 2000, 2002, 2004, 2005, 2006, 2007, 2009, 2010; Garcia-Domenech et al., 2008), which will be discussed and used for model purposes in this paper are defined by a family of formally similar algorithms. The very first graph-theoretical algorithm for these indices was devised by Randić in 1975 (Randić, 1975), and was, soon later, developed into a completely new chemical theory known as the molecular connectivity (MC) theory, by Kier and Hall (Kier & Hall, 1986, 1999). This theory has, since then, been

DOI: 10.4018/978-1-4666-4010-8.ch002

enriched and discussed in many ways (Pogliani, 2000; Garcia-Domenech et al., 2008; Trinajstić, 1992; Devillers & Balaban, 1999, Todeschini & Consonni, 2000; Diudea, 2000; Estrada, 2001; Li & Gutman, 2006). A central parameter of the molecular connectivity algorithms is the valence delta number, δ^v, which has recently undergone a radical transformation in the hands of Pogliani (Garcia-Domenech et al., 2008). Today this graph-theoretical number is defined in a way that it is able to encode not only the *sigma*-, *pi*- and non-bonding *n*-electrons but also the core electrons and the contribute of the suppressed hydrogen atoms, and it does it by the aid of concepts like general graphs or pseudographs and complete graphs. This important number has been defined in the following way,

$$\delta^v = \frac{(q + f_\delta^n)\delta^v(ps)}{(pr+1)} \tag{1}$$

Parameter $\delta^v(ps)$ is the vertex degree number of an atom in a pseudograph. Parameter q equals 1 or p; parameter $p \cdot r$ equals the sum of all vertex degrees in a complete graph, and it equals twice the number of its connections. Notice that parameter q has fixed values and it is not used as an optimizing parameter, something like Randić's variable index (Randić & Basak, 2001). The fact that q = 1 or p has the consequence that four possible sets of molecular connectivity indices can be obtained and they are: for $q = 1$ and $p = odd$ (1, 3, 5, ..) a K_p-(*p-odd*) set of indices is obtained, for $q = 1$ and p = sequential (1, 2, 3, …) a K_p-(*p-seq*) set is obtained, for $q = p$ and $p = odd$ a K_p-(*pp-odd*) set is obtained, and for $q = p$ and $p = seq$ a K_p-(*pp-seq*) set of indices is obtained. The rationale of this distinction resides in the fact that in many cases MC indices derived with *p* odd-valued show a superior model quality. Perturbation parameter f_δ takes care of the suppressed hydrogen atoms in a chemical graph or pseudograph, and is defined in the following way

$$f_\delta = [\delta^v_m(ps) - \delta^v(ps)]/\delta^v_m(ps) = 1 - \delta^v(ps)/\delta^v_m(ps) = n_H/\delta^v_m(ps) \tag{2}$$

Here, $\delta^v_m(ps)$ is the maximal $\delta^v(ps)$ value a heteroatom can have in a chemical HS-pseudograph when all bonded hydrogens are substituted by heteroatoms, and n_H equals the number of hydrogen atoms normally bonded to any heteroatom (C inclusive).

For completely substituted carbons (or heteroatoms): $n_H = 0$, $f_\delta = 0$ [as: $\delta^v_m(ps) = \delta^v(ps)$], while for alkanes δ and δ^v are now related as, $\delta^v(ps) = \delta$ and $\delta^v = (1 + f_\delta^n)\delta$ (in this case p = 1). Only for quaternary carbons (with $f_\delta = 0$) $\delta^v = \delta$ is obtained. Exponent n can have different values, and even negative values to allow for a strong perturbation value. The higher the value of n the lower the perturbation and for very high values ($n \geq 50$) the perturbation is practically zero. The reader should notice that the hydrogen perturbation allows to differentiate among *p*- and $\delta^v(ps)$-similar atoms, which differ in the number of bonded hydrogen atoms as in the case of the HS chemical pseudographs plus complete graphs of LiF, BeHF, BH_2F, and CH_3F.

For every n value a specific set of constant molecular connectivity index values is obtained, and this allows to check the amount of hydrogen perturbation, which is responsible for the best model. This parameter is not used as a variable parameter (even if it could) and this means that the number of molecular connectivity indices can be quite large as each set depends now on the values of q, p and n. These three values determine what could be defined a configuration of the molecular connectivity indices: K_p(*p-odd/seq*; *pp-odd/seq*)/f_δ^n.

Chemical Graphs, Pseudographs and Complete Graphs

Let us now spend two words about graphs, pseudographs (or general graphs) and complete graphs (Kier & Hall, 1986, 1999; Trinajstić, 1992; Devillers & Balaban, 1999, Todeschini &

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/graph-theoretical-indices-based-simple/77066

Related Content

A Hybrid Approach Based on Self-Organizing Neural Networks and the K-Nearest Neighbors Method to Study Molecular Similarity
Abdelmalek Amine, Zakaria Elberrichi, Michel Simonetand Ali Rahmouni (2013). *Methodologies and Applications for Chemoinformatics and Chemical Engineering (pp. 74-95).*
www.irma-international.org/chapter/hybrid-approach-based-self-organizing/77070

Hydrogen-Bonded Interactions in the Systems L-Cysteine - H2SeO3 and L-Cysteine -H2SeO4: A DFT Study
Alexei N. Pankratov, Nikolay A. Bychkovand Olga M. Tsivileva (2013). *Methodologies and Applications for Chemoinformatics and Chemical Engineering (pp. 169-179).*
www.irma-international.org/chapter/hydrogen-bonded-interactions-systems-cysteine/77076

Brain-like Processing and Classification of Chemical Data: An Approach Inspired by the Sense of Smell
Michael Schmukerand Gisbert Schneider (2011). *Chemoinformatics and Advanced Machine Learning Perspectives: Complex Computational Methods and Collaborative Techniques (pp. 289-303).*
www.irma-international.org/chapter/brain-like-processing-classification-chemical/45476

Nonlinear Partial Least Squares An Overview
Roman Rosipal (2011). *Chemoinformatics and Advanced Machine Learning Perspectives: Complex Computational Methods and Collaborative Techniques (pp. 169-189).*
www.irma-international.org/chapter/nonlinear-partial-least-squares-overview/45470

Structure –Activity Relationships by Autocorrelation Descriptors and Genetic Algorithms
Viviana Consonniand Roberto Todeschini (2011). *Chemoinformatics and Advanced Machine Learning Perspectives: Complex Computational Methods and Collaborative Techniques (pp. 60-94).*
www.irma-international.org/chapter/structure-activity-relationships-autocorrelation-descriptors/45466