Journal of Mathematical Chemistry Vol. 30, No. 4, November 2001 (© 2002)

Logical models of molecular shapes and their families *

A. Frolov [a], E. Jako [b] and P.G. Mezey [c,**]

[a] *Moscow Power Engineering Institute (Technical University), Chusovskaya, 11-7-8, Moscow 107203, Russia*
E-mail: frolovab@mpei.ru
[b] *Eötvös Loránd University, Department of Plant Taxonomy and Ecology, Tompa u. 17, Budapest 1211, Hungary*
E-mail: jakoeena@hotmail.com
[c] *Department of Chemistry and Department of Mathematics and Statistics, University of Saskatchewan, 110 Science Place, Saskatoon, SK, S7N 5C9 Canada*
E-mail: mezey@sask.usask.ca

Received 26 May 2001

Dedicated to the 80th birthday of Professor Frank Harary

A new discrete mathematical model of molecular shape is proposed, making use of the partition property of a representation of molecular shape. According to its geometrical and topological structure, a molecular surface can be partitioned into unbounded two-dimensional subsets (domains) and some common subsets of closures of two or more domains. The sets of these domains as a base of a finite topology, containing the Boolean n-cube as a lower Boolean sub-lattice of this topology, defines the domain of the proposed logical model. A logical function can be obtained that reflects the properties of the topological domains as well as the interrelations on the set of domains. Based on classical or quantum-chemical representations of molecular shape, these models allow one the implementation of methods of logical diagnostics in chemistry, and the definition of a metric on the set of molecular shape equivalence classes. The families of molecular shapes can be considered as sets of logical models. The proposed model is unified in the sense that the structures of differentiable and non-differentiable surfaces can be represented in the same mathematical framework. These logical models will also work for interpenetrations of the above types of surfaces.

KEY WORDS: molecular shape, logical models, Boolean n-cube, pattern recognition

0. Introduction

The study of molecular shapes and the recognition of similarity among molecules are of fundamental importance in the understanding of chemical and biochemical

* This article, dedicated to the 80th birthday of Professor Frank Harary, has been written as part of an Institute for Advanced Study, Collegium Budapest Fellowships scientific project: P.G. Mezey (1999–2000), E. Jako and A. Frolov (2000–2001).
** Corresponding author.

0259-9791/01/1100-0389/0 © 2002 Plenum Publishing Corporation

processes. Molecules with similar shape features often show similar chemical behavior on both the molecular and macroscopic levels. Molecular shape has a fundamental influence on both the static and dynamic properties of molecules; molecular interactions, molecular recognition, chemical reactions, and most other aspects of molecular behavior are directly dependent on molecular shape [1,2]. One may define and characterize different molecular bodies and surfaces depending on the selected geometrical or physical property one is interested in. For example, space filling models or alternative van der Waals surfaces (VDWSs), dot-body representations, molecular isodensity contours (MIDCOs), molecular electrostatic contours (MEPCOs), or cross-sections, representing intersections of any two of the above types of surfaces are among the possible choices [3–5].

At present, the most advanced shape representations of molecules can be obtained by combining the classical models with modern quantum-chemical approaches [6,7], whereas recent developments in molecular modeling and computational chemistry, have set the stage for the novel application of discrete mathematics to chemistry [8–11]. The known discrete mathematical models, serving as formal descriptors for the different chemical shape representations are based mainly on graph-theoretical [12–17], algebraic topological [18–21], and combinatorial [22–25] approaches, but continuous features may also be represented, by invoking a dependence of these models on continuous parameters. For example, a finite number of graph-theoretical or topological invariants are sufficient to characterize the molecular shape, that leads to a discrete characterization of the given family of continuous isoproperty contour surfaces (IPCOs) [26]. These invariants can be used to form shape codes, shape graphs or shape matrices, suitable for algorithmic shape comparisons. The division of the continuum of IPCOs into classes of shape groups with respect to the equivalence relation, allows us to consider only the essential shape features in conformational rearrangements of chemical processes. The corresponding algebraic-topological Shape-Group Methods (SGM) are described in detail in [1], and their various applications reviewed in [18–21].

In combinatorial shape models the equivalence relation for the evaluation of molecular similarity have been applied, based on recognition of square-cell configurations (“lattice animals”) or polycubes, as proposed by Harary and Mezey in [22–25,27,28]. These two- and three-dimensional models can be used for similarity analysis of different surfaces and molecular bodies on a natural, size-independent level of resolution. It should be emphasized, however, that a class of discrete mathematical models, based on logical functions, and implemented as logical models for the approximation of quantum-chemical and classical models of molecular shapes in chemical applications, is still undeveloped.

The purpose of this article is to introduce a novel *logical* approach and a class of logical shape models for chemical applications, based on the fundamental concept of logical *function.* Logical models of molecular shapes and their families provide implementation of methods of logical diagnostics in chemistry. As a known universal method for comparison of discrete objects, based on logical functional descriptions, there can be considered the method of Chegis and Yablonsky, originally proposed for testing switch-

ing circuits, described in [29]. This method was further developed and expanded to numerous applications in data analysis, including technical and logical diagnostics [30,31]. The above method can be implemented as soon as the logical functional descriptions of discrete objects or data are obtained. A direct implementation of this basic method to pattern recognition and clusterization theory, means considering the values of the logical functions (corresponding to some sets of variable values, possessing physical meaning) as the preliminary features of objects. The method allows one to chose a *test set*, that is, the minimum number of those features that preserve the separating property of the sample object models. Moreover, in terms of the initial object descriptions, it allows one the definition of the structural peculiarity that leads to the distinction of the equivalence classes of the modeling objects in terms of features from the given test set. Thus, by using the feature values from the test set, it is possible to calculate a diagnosis that identifies the whole object or equivalence classes of objects.

In other early contributions [32,33] elements of Boolean algebra were applied to solving problems of structural similarity and dissimilarity by using representations, called primitive relational structures.

A recently developed mathematical logical approach for design of logical networks, described by systems of Boolean functions, exploring separational properties of binary n-cube subsets, was proposed by Jako in [34]. This approach was first implemented for logical design [35] and classification of Boolean functions [34,36]. The corresponding algorithms for classification, exploring Disjunctive and Normal Forms were described in [37], whereas the Zhegalkin–Reed–Muller polynomial [38] and a novel analytical representation of Boolean and multiple valued logical functions, called as Jako Iterative Canonical Form (JICF) were implemented in algorithms proposed in [34,39,40]. For biological applications, the above mathematical logical approach was used in ecological studies for distinguishing spatial pattern in vegetation and species coalitions in plant communities [41]. On molecular level the same approach was used for similarity analysis and classification of macromolecules, considering their primary and secondary structures as classes of ordered sets [42]. The proposed methods [41,42] provide mapping of natural objects into binary n-cube, preserving partial ordering of structural components.

In the present article a class of novel logical models is introduced on the basis of known quantum-chemical and topological models of molecular shapes, and the Chegis–Yablonsky approach to data analysis and logical diagnostics, providing the tools for implementation of logical methods in the field of chemistry. The method proposed in this article, reflecting symmetrical relation of arbitrary degree on the modelled object structure, differs from logical methods [34–36,39–42], based on asymmetrical relations. Avoiding an explicit embedding of objects into a vector space, as well as their numerical characterization, it allows one a definition of a metric or semi-metric on the set of objects, providing an implementation of novel metric classification algorithms [43,44].

In the first section, the quantum-chemical nature of a molecular shape is discussed. The second section contains a detailed description of the proposed method of mathematical modeling of objects, assuming that the objects can be represented topologi-

cally. A topological representation, in this case, is the definition of the topological space (X, T), where X is the set of object points or elements, and T is a finite topology. It is assumed that the set X is partitioned into a finite number of subsets, some of which are chosen as a base of topology T. Some unary and n-ary ($n \geqslant 2$) symmetric relations on the chosen base are implemented to represent geometrical and topological properties of the set X. The main idea of the method is based on finite logical partitioning of the set X. A property of finite topology is used, where any finite topology contains a Boolean lattice with empty set. The Boolean cube isomorphic to this lattice can be considered as a field of definition (domain) of a logical function, reflecting the previously mentioned properties and relations. The proposed model can be implemented for advanced similarity and dissimilarity analysis of various n-dimensional ($n \geqslant 2$) objects on different organizational levels, such as molecular shapes and their cross-sections, multi-dimensional energy hypersurfaces, [46] or discrete N-dimensional systems for ecological studies. The third section is devoted to logical models of molecular shapes and their families. The connection with the second section arises in that the topology of a shape is based on a partition of surfaces into 2-, 1-, and 0-dimensional subsets with some finite sets of properties. These subsets interrelate with each other through known symmetric relations. This allows for the implementation of the proposed method of logical modeling. Characteristic properties of various types of models, such as the quantum-chemical models based on IPCOs, the classical models based on van der Waals surfaces, and models based on intersections or superpositions of the classical and quantum-chemical model surfaces are discussed. It is shown that proposed models possess the same separable properties as known models, based on shape code matrices or VDW graphs [1,9,10]. Examples are given, illustrating how the proposed models can be used for calculation of structural peculiarities for distinction of equivalence classes of molecular shapes. Generalization and application possibilities of functional models for decision making systems are discussed in concluding remarks.

1. The quantum-chemical and topological nature of molecular shapes

Modern quantum chemical and topological methods allow one to study many molecular problems where experimental information is not available. The quantum chemical representation of molecular shape is based on the concepts of nuclear configuration and chemical bonding. Both of these concepts can be reformulated in terms of topology. The basic idea of three-dimensional topological approach, formulated first by P.G. Mezey and known as the GSTE principle, is that by using adequate topological descriptions for molecular shapes, geometrical similarity is treated as topological equivalence.

For further analysis it is appropriate to consider the shape description, based on different physical properties, and/or similar geometrical construction within a common framework. For any physical molecular property P that is described by a 3D function $P(r)$ which is continuous in r, such as the electronic density, the electrostatic potential $V(r)$ or composite nuclear potential $Vn(r)$, the level sets $F(a)$ for any constant value a of function $P(r)$ can be defined as the following collection of points:

$$F(a) = \{\mathbf{r}\colon\ P(\mathbf{r}) < a\}. \tag{1.1}$$

The boundary surfaces $G(a)$ of level sets $F(a)$ are the isoproperty contours (IPCOs) defined as

$$G(a) = \{\mathbf{r}\colon\ P(\mathbf{r}) = a\}. \tag{1.2}$$

The entire 3D property function $P(r)$ can be represented by an infinite family of IPCOs, by taking one such surface $G(a)$ for each value of the contour parameter a, throughout the entire range $amin \leqslant a \leqslant amax$. The minimum and maximum values *amin* and *amax* of the contour threshold a depend on the property P and the actual molecular system studied.

It is possible to generate a function of density-dependent radius $r_A(a)$ for a spherical representation of each atom A, and construct a fused sphere VDWS representation of a molecule that mimics a boundary surface $G(a)$ of an IPCO for any selected density value a. Therefore, the fused spheres model provides an approximation of the molecule for a given threshold value a. It is is important to note however, that the techniques developed for the shape characterization of smooth, differentiable contour surfaces (such as MIDCOs, MEPCOs, or molecular orbitals) are not always applicable to all molecular surfaces, like VDWS, solvent accessible surfaces or union surfaces, defined by some other criteria [46,48]. For example, at every point r of a VDWS where two or more atomic spheres interpenetrate one another, the surface is not smooth and is not differentiable. In order to perform the shape analysis of VDWS one needs the determination of the arcs on the surface resulting from the interpenetration of the atomic spheres, as well as the knowledge of those points that lie on the surface of three or more spheres. The above information can be obtained from the available programs used to construct the VDWSs, or it can be deduced from a graphical display.

In the present paper we will consider models for description of molecular shapes, based on quantum-chemical representations using molecular isoproperty contour surfaces (or IPCOs), and classical hard sphere representations using alternative VDWSs as well as cross-sections and superpositions of the above types of molecular surfaces.

The VDWS for an N-atomic molecule can be defined by two sets:

$$C_x = \{x_1, x_2, \ldots, x_N\}, \quad x_i \in \mathbb{R}^3, \tag{1.3a}$$

and

$$C_p = \{p_1, p_2, \ldots, p_N\}, \quad p_i \in \mathbb{R}, \tag{1.3b}$$

where the set C_x defines the position vectors for the nuclei, and the set C_p contains the atomic van der Waals radii, corresponding to the atoms [5,9,10].

Within a global approach to the study of molecular deformations and their relations to molecular identity, it is advantageous to use the concept of nuclear configurations. Furthermore, it is appropriate to restrict our study considering only the internal nuclear configuration (i.e., the relative arrangement of the nuclei with respect to one another, when a molecule far removed from other molecules and from sources of external field). Each internal configuration K for a molecule contains N, $N \geqslant 4$, nuclei, which can be

defined as a $(3N-6)$-dimensional vector. This vector belongs to the metric space M of internal configurations.

The goal of further analysis is to obtain a hierarchy of surfaces (Gn) for some given nuclear configuration K, obtained by partitioning the surface $G(K)$.

In metric space M for each point a of the VDWS surface $G(K)$, there can be defined corresponding distances $d(a, x_i)$, $i = 1, \ldots, N$. Therefore, the surface $G(K)$ is partitioned into subsets

$$D_k = \{a \mid a \in G(K)\}, \tag{1.4}$$

where k is the number of indices I such that $d(a, x_i) = i$, i.e., k is the number of spheres to which the given point a belongs. The sets D_k can be thought as union of their maximal connected components D_k^j, $j = 1, \ldots, m_k$:

$$D_k = \bigcup_{j=1}^{m_k} D_k^j. \tag{1.5}$$

The corresponding formally defined surfaces in 3D metric space can be connected or disconnected, but for an isolated molecule they are bounded. The shapes can be easy mapped into usual Cartesian 3D space. The most important property of these approximations is that the corresponding surfaces can be partitioned into well-defined subsets according to their geometrical and topological properties. This property allows us to consider a finite set of shapes of a given geometry and topology instead of infinite set of shapes corresponding to distinct nuclear configurations K and distinct thresholds a. In other words, it allows one to define an equivalence relation on the Cartesian product $M \times A$, where A is the set of possible thresholds

$$(K_1, a_1) \approx (K_2, a_2),$$

if IPCO's or VDWS's for (K_1, a_1) and for (K_2, a_2) satisfy the same topology. It means, that each class can be represented by a unique pair, and a unique shape can be considered for the whole class.

Now, it is possible to consider families of equivalence classes corresponding to some given electron configuration K with different threshold values a

$$a_1 < a_2 < \cdots < a_k,$$

or to families of equivalence classes corresponding to some given threshold value a and different electron configurations $K_1, K_2, \ldots, K_t$, according to the dynamics of a molecule's behaviour (or its trajectory in the nuclear configuration space M). In the same manner, curvature parameter b [1] can be taken into account.

We assume that equivalence classes consist of surfaces partitioned on the same number of subsets, possessing the same geometrical properties and participating in the same symmetrical relations. This allows us to propose a new unified method with respect to distinct quantum chemical, classical and mixed model logical method that also

allows us to distinguish shapes of distinct classes as well as obtain minimal description of difference and introduce a metric on the set of molecular shape equivalence classes.

2. Logical modeling and comparison of objects of geometrical or discrete structure

2.1. Finite topology property

In this section, a logical model for objects, which have geometries or discrete structures that allow for the partitioning into a finite number of pairwise, non-intersecting parts, is proposed. The logical model reflects properties of these parts, as well as their interrelations. The creation of the model is based on the concepts of topology and topological space, as well as on the concept of logical function [49–54]. Recall that a *topological space* is a pair (X, T) that consists of a nonempty set X and a subset T of $P(X)$ such that $\emptyset$ and X belong to T, where $P(X)$ is the set of all the subsets of the set X. Moreover, T is closed with respect to unions and finite intersections. Such a set T is called a *topology*. In particular, any finite lattice T of subsets is a topology if $\emptyset, X \in T$. The elements of T are called *open sets*. The set $M \in X$ is *closed* (*open*), if its complement M' is an open (closed) set. Sets that are both open and closed are called *clopen* sets. For example, the empty set ($\emptyset$) and the set X are clopen sets.

For the purposes of this article, only finite topologies are used. Each finite topology can be defined by the set of their atoms, i.e., nonempty subsets $a \in T$, such that $\forall b \in T$, $b \neq a$, $b \neq \emptyset$, $b \not\subset a$. This set of atoms forms the topology base B where each element of T, except $\emptyset$ and X, is a union of some sets in B. Moreover, a finite topology contains all the sets that are unions of some sets in B. In this case, topology T is called a *topology of base* B. These unions, together with the empty set $\emptyset$, form the Boolean sublattice. Topology T contains all the elements of this Boolean sublattice and the set X. Furthermore, T is a distributive lattice. Consequently, it can be noted that topology T has two Boolean sublattices: the previously mentioned sublattice, and the sublattice $\{U, X\}$, where U is the upper bound of the former sublattice. In the case of the coincidence $U = X$, the latter sublattice degenerates and becomes a trivial Boolean sublattice containing only one element, X.

Example 2.1. Consider the Cartesian product $X = \mathbb{R} \times \mathbb{R}$, and Jordan curve $J \in X$; let A_1 and A_2 be two subsets of X truncated at the points P_1 and P_2, and separated by curve J. Then

$$X = A_1 \cup A_2 \cup J \cup \{P_1\} \cup \{P_1\}, \left\{A_1, A_2, J, \{P_1\}, \{P_2\}\right\}$$

is considered a partition of set X (figure 1).

Consider the topological space $\{X, T\}$ with topology T of base $\{J, \{P_1\}, \{P_2\}\}$. This topology contains all the bounded subsets of set X (the open sets of topology) as well as the set X, and the empty set $\emptyset$ (figure 2).

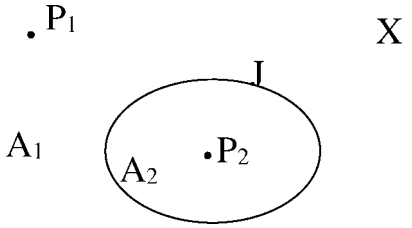

Figure 1. The partition $\{A_1, A_2, J, \{P_1\}, \{P_2\}\}$ of the Cartesian product $X = \mathbb{R} \times \mathbb{R}$.

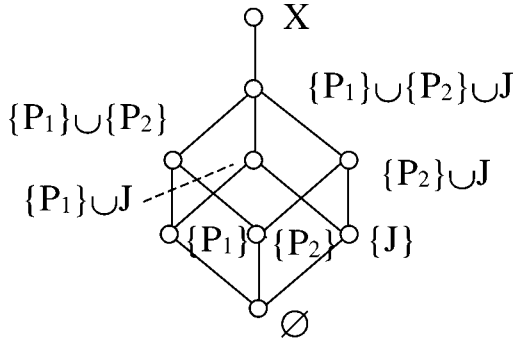

Figure 2. The topology of base $\{J, \{P_1\}, \{P_2\}\}$. There are two Boolean sublattices $\{\emptyset, \{P_1\}, \{P_2\}, \{P_1\} \cup \{P_2\}, \{P_1\} \cup J, \{P_2\} \cup J, \{P_1\} \cup \{P_2\} \cup J\}$ and $\{\{P_1\} \cup \{P_2\} \cup J, X\}$.

There are two Boolean sublattices:

$$\big\{\emptyset, \{P_1\}, \{P_1\}, \{P_1\} \cup \{P_2\}, \{P_1\} \cup J, \{P_2\} \cup J, \{P_1\} \cup \{P_2\} \cup J\big\}$$

and

$$\big\{\{P_1\} \cup \{P_2\} \cup J, X\big\}.$$

Example 2.2. Consider the set X and its partition as shown in figure 1. The topological space $\{X, T\}$ can be constructed with topology T of base $\{A_1, A_2\}$. This topology contains all the unbounded subsets of the set X (open sets of topology), as well as the set X and the empty set $\emptyset$. There are two Boolean sublattices: $\{\emptyset, A_1, A_2, A_1 \cup A_2\}$ and $\{A_1 \cup A_2, X\}$ (figure 3).

These examples illustrate that the definition of a topology allows for the choice of closed sets of a topology from bounded subsets of the set X, as well as from unbounded sets. A mixed variant is possible as well. For example, a topology of base $\{A_1, A_2, J\}$ can be chosen, or the topology based on the whole partition $\{A_1, A_2, J, \{P_1\}, \{P_2\}\}$ of the set X can be chosen. In the latter case, all the elements of topology T are clopen sets, and topology T is a Boolean lattice where the upper bound U coincides with the set X.

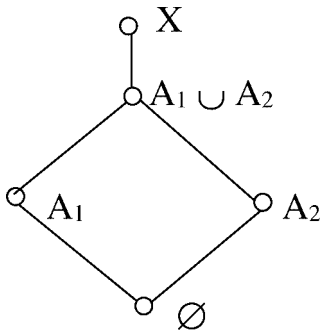

Figure 3. The topology of base $\{A_1, A_2\}$. There are two Boolean sublattices $\{\emptyset, A_1, A_2, A_1 \cup A_2\}$ and $\{A_1 \cup A_2, X\}$.

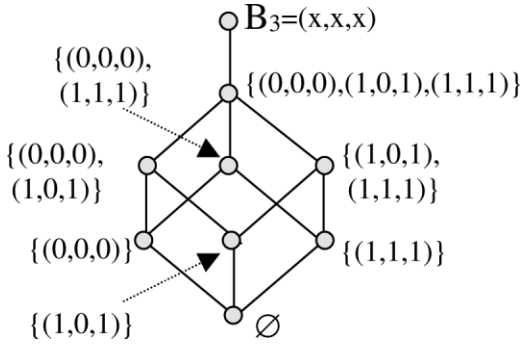

Figure 4. Finite topological space $\{B_3, T\}$ with topology T, based on the sets $\{(0, 0, 0), (1, 0, 1), (1, 1, 1)\}$. $B_3 = (x, x, x) = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 0), (1, 1, 1)\}$.

The set X may be a continuous set, as in the previous examples, or it may be a discrete set. In particular, it may be finite. For example, consider the Boolean n-cube B_n, $n = 3$, as a set X and the partition $\Pi = \{\{(0, 0, 0), (1, 0, 0)\}, \{(0, 1, 0)\}, \{(0, 0, 1)\}, \{(1, 1, 0)\}, \{(1, 0, 1)\}, \{(0, 1, 1)\}, \{(1, 1, 1)\}\}$ that consists of all one-element subsets of the set X. From this set, some subset of Π can be chosen as a base of topology T. Figure 4 represents a particular case of topological space (B_3, T), with topology T, based on the sets $\{(0, 0, 0)\}, \{(1, 0, 1)\}, \{(1, 1, 1)\}$. On the other hand, the topology based on Π contains $2^8 = 256$ clopen subsets. It is a Boolean lattice isomorphic to the lattice $P(X)$.

Generalizing the properties of finite topologies, it can be noted that, considering any set X and its finite partition $\Pi = \{A_1, A_2, \ldots, A_n\}$, one can construct a topological space (X, T) with finite topology T, based on some subset of Π. The topology contains two Boolean sublattices: an *upper lattice* $\{U, X\}$ and a *lower lattice* $T \setminus \{X\}$. In the case of the base set $B = \Pi$, both Boolean lattices coincide and the upper Boolean sublattice degenerates and becomes a trivial Boolean lattice containing only one element. This finite topology property is depicted in figure 5.

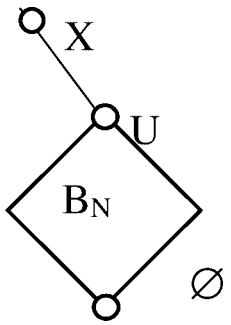

Figure 5. The finite topology property: considering any set X and its finite partition $\Pi = \{A_1, A_2, \ldots, A_n\}$, one can construct a topological space (X, T), with finite topology T, based on some subset of Π. The topology contains two Boolean sublattices: an *upper lattice* $\{U, X\}$ and a *lower lattice* $T \setminus \{X\}$.

2.2. Logical models of topological spaces with finite topology

Now, let us show how to construct logical models of a topological space of a finite topology taking into account properties of base sets as well as their interrelations. Suppose that each base set has one of the finite number properties $\{0, 1, 2, \ldots, K-1\}$. Moreover, assume, that the base sets participate in the symmetric relations

$$\rho_0^{m_1}, \ldots, \rho_i^{m_i}, \ldots, \rho_{S-1}^{m_{s-1}}$$

where a relation, $\rho_i^{m_i}$, is a symmetric relation if, together with any m_i ordered set (m_i-tuple), it contains any of its permutations. (We assume that the system of these relations is orthogonal, i.e., each m_i-tuple participate in only one m_i-dimensional relation.) In addition, it should be noted that the mentioned properties and relations should be defined based on the real properties of the geometrical or discrete structures of the modelled objects. Also, the properties of all the elements from the base set of the topology, as well as the relations on the set of base elements, are numbered $0, \ldots, K-1$ and $0, 1, \ldots, S-1$, respectively.

The information about geometry, as well as about the properties and interrelations of the elements from the base of topology can be represented as a logical function $f(x_1, \ldots, x_N)$ or as an equivalent system of Boolean functions.

We use the Boolean N-cube B_N as a field of definition (domain) of a logical function $f(x_1, \ldots, x_N)$, $f : B_N \to \{0, 1, \ldots, k-1\}$, $k = \max(K, S)$. This function maps the elements from the base of the topology to the values defining their properties from the set $\{0, 1, \ldots, K-1\}$ of numbers of the properties. Later, if the sets $A_1, \ldots, A_t$ from the base of the topology belong to relation

$$\rho_t^{mt},$$

then the function takes the topology element corresponding to the union of these base sets to the value t from the set $\{0, 1, \ldots, S-1\}$. Finally, the function f takes all the remaining elements of Boolean N-cube B_N to the zero value.

Example 2.3. Let (X, T) be a topological space, where $X = \mathbb{R} \times \mathbb{R}$, T is topology, based on the partition Π (figure 1). Assume that the zero-dimensional subsets and the one-dimensional subsets possess properties 0 and 1, whereas the two-dimensional subsets A_2 and A_1 have the properties 2 (internal) and 3 (external), respectively. Moreover, a symmetrical binary relation ρ_1^1 is defined on the set Π:

$$\rho_1^1 = \big\{(B_i, B_j) \mid B_i \text{ and } B_j \text{ have the common boundary, or } B_i \text{ is in boundary of } B_j, \text{ or } B_j \text{ is in boundary of } B_i\big\}.$$

The 3-valued function $f(x_1, x_2, x_3, x_4, x_5)$ of variables x_1, x_2, x_3, x_4, x_5 corresponding to the sets $A_1, A_2, J, \{P_1\}, \{P_2\}$ can be defined as follows:

$$\begin{aligned}
&f(1,0,0,0,0) = 3 \quad (A_1 \text{ is an external two-dimensional set}),\\
&f(0,1,0,0,0) = 2 \quad (A_2 \text{ is an internal two-dimensional set}),\\
&f(0,0,1,0,0) = 1 \quad (J \text{ is an internal one-dimensional set}),\\
&f(1,1,0,0,0) = 1 \quad (A_1 \text{ and } A_2 \text{ have the common boundary}),\\
&f(1,0,1,0,0) = 1 \quad (J \text{ is in boundary of } A_1),\\
&f(0,1,1,0,0) = 1 \quad (J \text{ is in boundary of } A_2),\\
&f(1,0,0,1,0) = 1 \quad (P_1 \text{ is in boundary of } A_1),\\
&f(0,1,0,0,1) = 1 \quad (P_2 \text{ is in boundary of } A_2).
\end{aligned}$$

The function f takes the remaining binary 5-tuples to the value 0, in particular

$$f(0,0,0,1,0) = f(0,0,0,0,1) = 0 \quad (P_1 \text{ and } P_2 \text{ are zero-dimensional subsets}).$$

This function contains all the information required for the reconstruction of the topological space, and in particular, it allows figure 1 to be redrawn.

If we construct analogous model for the other partition of the same set X shown in figure 6, we obtain other logical function $f(x_1, x_2, x_3, x_4, x_5)$ that differs from the function $f(x_1, x_2, x_3, x_4, x_5)$ in the following positions:

$$f'(1,0,0,1,0) = 0, \quad f'(0,1,0,1,0) = 1$$

(point P_1 is replaced from boundary of the set A_1into the boundary of the set A_2).

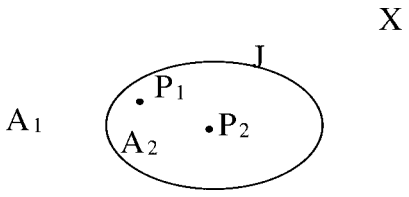

Figure 6. The other partition $\{A_1, A_2, J, \{P_1\}, \{P_2\}\}$ of the Cartesian product $X = \mathbb{R} \times \mathbb{R}$.

3. Logical models of molecular shapes

This section is devoted to logical models of molecular shapes and their families. It relates to the previous section in that the topology of a shape is based on the partitioning of a surface on two-, one-, and zero-dimensional subsets with finite set properties. These subsets interrelate with each other through known relations. This allows implementation of the finite topology principle, based on the method of logical modeling introduced in the previous section.

3.1. Logical models of isoproperty contour surfaces

In fact, for the characterization of the shapes of molecular contour surfaces, such as MIDCOs and MEPCOs, or an interpenetration of both, the surface is subdivided into domains satisfying some local shape criteria. This is discussed in detail in [1, chapter 5]. These absolute or relative shape domains satisfy some geometrical or physical properties. For example, these domains may be thought as unbounded locally convex, locally concave, or locally-saddle-type subsets of a surface or as unbounded sets deleted from the shape. More exactly, domains D can be truncated by subdivision into three sets: the unbounded set C where $C \subset D$, the unbounded set $D' \subset D$, and the common boundary J of the sets C and D, i.e., $J = \mathrm{clos}(C) \cap \mathrm{clos}(D)$, where clos is the set-theoretical operation of closure. Set C is referred to as a *deleted* set and set D' is referred to as a *truncated* set.

If all the truncated domains, non-truncated domains, deleted sets, and one-dimensional boundaries (with the possible addition of the zero-dimensional boundaries of boundaries) are taken into account, a partition of the surface can be obtained. This partition satisfies all the conditions required by the proposed logical modeling method: by considering the molecular surface before deletions as a set X, the partitions act as subsets of X. Depending on the goal of modeling, an appropriate subset of these subsets can be chosen as a base for topology T. For example, all these subsets can be chosen as a base, and in general, choosing all of the domains and deleted sets as a base proves to be most useful. Yet, regardless of which base is chosen, the elements of the topology base interrelate and possess some finite number properties. Finally, it should be noted that the same arguments apply to the adequateness of logical modeling in cases if IPCO's, VDWS's, or interpenetrating of the above types of surfaces are used.

Consider some common properties of lower Boolean lattices under the assumption that all the nontruncated domains, truncated domains, as well as all the deleted sets are used as a base sets of topology T. For clarity, all of these sets are henceforth called domains and are denoted as $D_1^{i_1}, \ldots, D_1^{i_N}$.

In the general case, $T_s^{i_1,\ldots,i_s}$ (for $s \geqslant 2$) denotes the maximal finite set that contains $|T_s^{i_1,\ldots,i_s}|$ nonempty subsets of maximal connected components which are subsets of the intersection of the domains $D_1^{i_1}, \ldots, D_1^{i_s}$ closures. It is assumed that each element of the set $T_s^{i_1,\ldots,i_s}$ is obtained from one of such maximal connected component, after deleting all the points that belong to closure of at least one other domain. If the deletion of such points leads to an empty set, that should not be represented in the set $T_s^{i_1,\ldots,i_s}$.

Assuming that $T_1^{i_1} = \{D_{i_1}\}$, the sets $T_s^{i_1,\dots,i_s}$ correspond to elements $D_1^{i_1} \cup \cdots \cup D_1^{i_s}$, $s = 1, \dots, N$, of a lower Boolean sublattice. By taking into account the previously mentioned isomorphism φ, the sets correspond to the elements of Boolean N-cube as well. Therefore, the properties of the sets $T_s^{i_1,\dots,i_s}$ can be used to define a logical function f, as was introduced in the previous section. In actuality, the final definition depends on a preliminary geometrical topological model of a molecular shape, but it is useful to note some common agreements. For example:

$$f(\underbrace{0, \dots, 0,}_{i-1} 1, \underbrace{0, \dots, 0}_{N-i}) = 0, \tag{3.1}$$

if D_1 is a deleted set.

Moreover, $f(a_1, \dots, a_i, \dots, a_N) = 0$, if $(T_s^{i_1,\dots,i_s}) = \emptyset$, where i_j are the numbers of the binary elements $a_{i_j} = 1$.

Furthermore, for MIDCO- or MEPCO-based molecular shapes, the function is defined such that

$$f(\underbrace{0, \dots, 0,}_{i-1} 1, \underbrace{0, \dots, 0}_{N-i}) = 1 + \mu(D_i), \tag{3.2}$$

and

$$f(\underbrace{0, \dots, 0}_{i-1}, 1, \underbrace{0, \dots, 0}_{j-i-1}, 1, \underbrace{0, \dots, 0}_{n-i-j}) = \begin{cases} 1, & \text{if } \operatorname{clos}(D_i) \cap \operatorname{clos}(D_j) = T_2^{i,j} \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \tag{3.3}$$

where $\mu(D_i)$ is the curvature index of the domain D_i.

Example 3.1. Consider the molecular shape partition, [1] shown in figure 7. It includes one locally saddle-type domain D_1^1, three locally convex domains D_1^2, D_1^3, D_1^4, and one locally concave domain D_1^5. There are four boundaries (closed lines)

$$L^6 = D_2^{1,2}, \quad L^7 = D_2^{1,3}, \quad L^8 = D_2^{1,4} \quad \text{and} \quad L^9 = D_2^{1,5}.$$

The union of these eight sets coincides with the shape of the molecule, as the sets do not intersect mutually. Therefore the set

$$\left\{D_1^1, D_1^2, D_1^3, D_1^4, D_1^5, D_2^{1,2}, D_2^{1,3}, D_2^{1,4}, D_2^{1,5}\right\}$$

is a partition of the molecular surface. If the topological base

$$\left(D_1^1, D_1^2, D_1^3, D_1^4, D_1^5\right),$$

is chosen, then a finite topology that contains a Boolean sublattice isomorphic to the Boolean cube B_5 is obtained. A shape matrix $s(a, b) = s(a, 0)$ corresponding to some

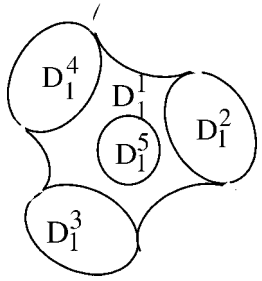

Figure 7. The shape domains of a MIDCO surface $G(a)$ relative to some curvature parameter b. The locally saddle-type D_1^1, locally convex D_1^2, D_1^3, D_1^4, and locally concave D_1^5 domains, as well as their pairwise common boundaries (cyclic curves), are shown [1].

isodensity threshold a and a curvature parameter b of zero describes the shape domains of ordinary local convexity of the MIDCO surface in figure 7 as follows:

$$s(a,b) = \begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{matrix}.$$

The diagonal elements $s(a,b)_{i,i}$, $i = 1, 2, 3, 4, 5$, are equal to the common curvature index $\mu(a,b)$ for the points within the domain D_1^i, while the off-diagonal elements $s(a,b)_{i,j}$ are equal to 1 if the domains D_1^i and D_1^j have a common boundary that defines them as being "neighboring domains", while the off diagonal elements equal 0 if there is no such common boundary between the domains.

According to the proposed logical modeling method, the logical model can be represented as a 4-valued function $f(x_1, x_2, x_3, x_4, x_5)$ of five variables. The variables x_1, x_2, x_3, x_4, and x_5 correspond to the domains

$$D_1^1, D_1^2, D_1^3, D_1^4, D_1^5.$$

The values of this function on binary 5-tuples, which correspond to the domains are the following:

$f(1,0,0,0,0) = 2$,

(rule (3.2), D_1^1 is a locally saddle-type domain, with $\mu(D_i) = 1$);

$f(0,1,0,0,0) = f(0,0,1,0,0) = f(0,0,0,1,0) = 3$,

(rule (3.2), D_1^2, D_1^3, and D_1^4 are locally convex domains, with $\mu(D_i) = 2$);

$f(0,0,0,0,1) = 1$,

(rule (3.2), D_1^5 is a locally concave domain, with $\mu(D_i) = 0$);

$f(1,1,0,0,0) = f(1,0,1,0,0) = f(1,0,0,1,0) = f(1,0,0,0,1) = 1$,

(rule (3.3), corresponds to the nonempty sets $D_2^{1,2}$, $D_2^{1,3}$, $D_2^{1,4}$, $D_2^{1,5}$ which are the boundaries between the domains).

The function assigns a value of zero to all of the remaining 5-tuples.

To summarize, it has been shown that the shape matrix describing MIDCOs or MEPCOs can be used in the construction of a logical model of the MIDCOs or MEPCOs, expressed as a logical function $f : B_N \to \{0, 1, 2, 3\}$ and, hence, as a system of Boolean functions. The function f is referred to hereafter as a *molecular contour function* (MCOF).

Note, that representation of molecular shapes by MCOF as well as by shape matrix may be ambiguous, depending on linear ordering of two-dimensional domains of a molecular shape. Nevertheless, the following statement is valid.

Statement 3.1. There exists a one-to-one correspondence between the set of molecular contour functions $f(x_1, \ldots, x_i, \ldots, x_n)$ representing distinct MIDCOs and the set of shape matrices.

Indeed, for a given ordering of two-dimensional domains the shape matrix can be derived from the MCOF, and the MCOF can be obtained from the shape matrix.

3.2. Logical models of van der Waals surfaces

The considered logical models can be implemented for VDWSs only in the special case when the sets

$$T_s^{i_1,\ldots,i_s}$$

in the case where $s = 2$, contain at most one set, or, in the cases where $s > 2$, are empty sets.

Therefore, the logical modeling of VDWSs is more complex than in the case of the MIDCO-based counterparts. Two items need to be accounted for:

(a) the sets

$$T_2^{i_1,i_2}, \quad s = 2,$$

can contain more than one set; and

(b) the sets

$$T_s^{i_1,\ldots,i_s}, \quad s > 2,$$

can be nonempty. Note that these sets can consist of more than one, but less than N, maximal connected components, where N is the number of nuclei. All elements of the sets

$$T_s^{i_1,\ldots,i_s}, \quad s > 2,$$

are either single-membered sets or empty sets. The single-membered sets consist of a unique element that belongs to a sphere or spheres represented on the VDWS by some domains

$$D_1^{i_1}, \ldots, D_1^{i_s}.$$

If the number of spheres to which an element belongs is denoted as k, then, in the general case, s does not equal k.

Although the theory of logical modelling can be generalized, in the present discussion for clarity it will be restricted to the case where shapes satisfy the criterion that a *VDWS contains at least one domain from each sphere*. In this case elementary geometry shows that, in the case of $s > 2$, the sets $T_s^{i_1,\dots,i_s}$ satisfy the following properties:

(a) $|T_s^{i_1,\dots,i_s}| \leqslant 2$;

(b) $|T_s^{i_1,\dots,i_s}| = 2$ implies the equality $k = s$ for elements of single-membered sets from $T_s^{i_1,\dots,i_s}$;

(c) if $|T_s^{i_1,\dots,i_s}| = 1$ when $s > 2$ then $k \in \{s, s+1\}$ for an element of a single-membered set from $T_s^{i_1,\dots,i_s}$.

According to the proposed logical modeling method, the construction of logical models of molecular shapes represented by VDWSs can be achieved by utilizing the following rules:

- The topological space (VDWS, T), where VDWS is the modeled van der Waals surface, and T is a finite topology based on the set B of the VDWS domains

$$D_1^j, \quad j = 1, \dots, N$$

should be chosen. It is assumed that some of these domains have been deleted, while others have been left in the VDWS.

- The Boolean function $f : B_{m_1} \to \{0, 1, \dots, N\}$ is defined as follows:

$$f(\underbrace{0, \dots, 0}_{i-1}, 1, \underbrace{0, \dots, 0}_{N-i}) = \begin{cases} 0 & \text{if } D_i \text{ is a deleted domain,} \\ 3 & \text{otherwise,} \end{cases}$$

according to the rules (3.1) and (3.2);

$$f(\underbrace{0, \dots, 0}_{i-1}, 1, \underbrace{0, \dots, 0}_{j-1}, 1, \underbrace{0, \dots, 0}_{N-i-j}) = \left|\left(T_2^{i,j}\right)\right|, \tag{3.6}$$

in the case, where the weight (the number of non-zero elements) of the N-tuple $r(a_1, \dots, a_n)$, is greater than two, the indices $i_1, \dots, i_s$, $s > 2$, correspond to the position of the non-zero elements of the binary N-tuple

$$f(a_1, \dots, a_n) = \begin{cases} 0 & \text{if } \left|\left(T_s^{i_1,\dots,i_s}\right)'\right| = 0, \\ 1 & \text{if } \left|\left(T_s^{i_1,\dots,i_s}\right)'\right| = 1, \text{ and } s = k, \\ 3 & \text{if } \left|\left(T_s^{i_1,\dots,i_s}\right)'\right| = 1, \text{ and } s \neq k, \\ 2 & \text{if } \left|\left(T_s^{i_1,\dots,i_s}\right)'\right| = 2. \end{cases} \tag{3.7}$$

As previously stated, k is the number of spheres to which a unique element of the set $(T_s^{i_1,\dots,i_s})'$ belongs.

Note that the function f takes N-tuples of weight 1 to values that are the same as for MIDCOs and MEPCOs. Therefore, the modeling method may be implemented upon models such as VDWSs when treated as a form of interpenetrating molecular surfaces.

The domain of f coincides with the N-cube isomorphic to the lower Boolean sublattice of the finite topology T. Assuming that the VDWS contains at least one domain from each sphere, it can be shown that the function $f(x_1, \ldots, x_n)$ is well defined. This logical function can be called the *VDWS function*.

A VDW graph can be derived utilizing the VDWS function. If the VDWS function $f(a_1, \ldots, a_i, \ldots, a_n)$ is represented as a marked binary N-cube, where the vertices are marked by the values of the logical function, the VDW graph can be created by implementing the following algorithm:

1. Delete all the vertices of weight two or greater labeled zero, as well as the incident edges to and from these vertices,
2. Connect any isolated vertices $(a_1, \ldots, a_i, \ldots, a_N)$ of weight greater than two with all the vertices $(a'_1, \ldots, a'_i, \ldots, a'_N)$ of weight two such that $(a'_1, \ldots, a'_i, \ldots, a'_N) \leqslant (a_1, \ldots, a_i, \ldots, a_N)$,
3. Split the vertices (and their incident edges) $(a_1, \ldots, a_i, \ldots, a_N)$ of weight two or greater such that $f(a_1, \ldots, a_i, \ldots, a_N) = 2$. The splitting should be done $f(a_1, \ldots, a_i, \ldots, a_N)$ times,
4. Replace vertices $(a_1, \ldots, a_i, \ldots, a_N)$ on the $(r + 1)$th level of the graph of weight greater than two that are given the label 3 by the function.

By reversing the order of the rules, a graphical representation of the VDWS function can be obtained from the VDW graph. Therefore, both the VDWS function and the VDW graph represent the same VDWS and the following statement is valid (although representation of molecular shapes by VDW graph as well as by MCOF is ambiguous in defense on linear ordering of two-dimensional domains of a molecular shape).

Statement 3.2. There exists a one-to-one correspondence between the set of VDWS functions representing distinct VDWSs and the set of VDW graphs.

Example 3.2. As a simple example, consider an AB_2 type molecule with oscillation dynamics consisting of three equivalence classes of shapes, whose representative geometrical topologies are shown in figure 8. Graphical representations of the VDWS functions for the VDWSs in figures 8(a)–(c) are given in figures 9(a)–(c). They are obtained according to the rules (3.2), (3.3), (3.6), (3.7). Note that the value $f_c(1, 1, 1) = 2$ because there are two single-membered components in the intersection of three closures

$$\bigcap_{i=1,2,3} \mathrm{clos}(D_1^i) \tag{3.8}$$

of domains in figure 8(c) (rule (3.6), $s = k$). The value $f_b(1, 1, 1) = 1$ because there is only one single-membered component in the intersection (3.8) of three closures of domains in figure 8(b) (rule (3.6), $s = k$).

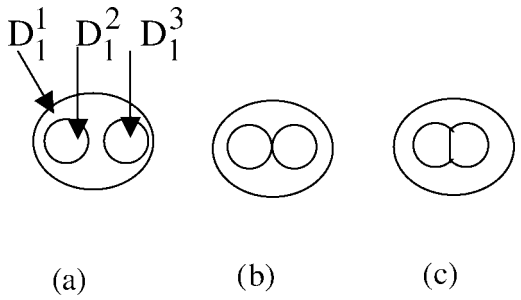

Figure 8. The topologically different VDWSs in the bending oscillation corresponding to the "umbrella inversion" of an AB_2 type molecular system.

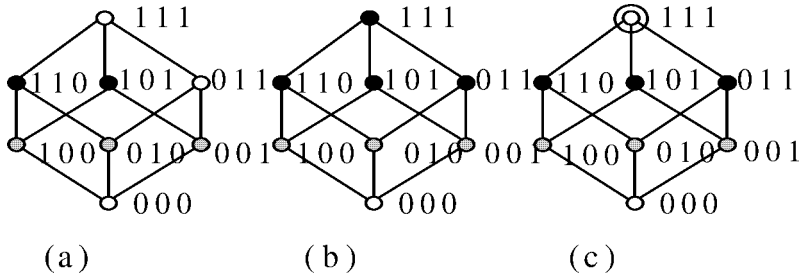

Figure 9. Structural representations of VDWS functions in the bending oscillation corresponding to the "umbrella inversion" of an AB_2 type molecular system (a, b). The white, black, double circled and gray points correspond to $N = 0, 1, 2$ and 3 values of the VDWS functions.

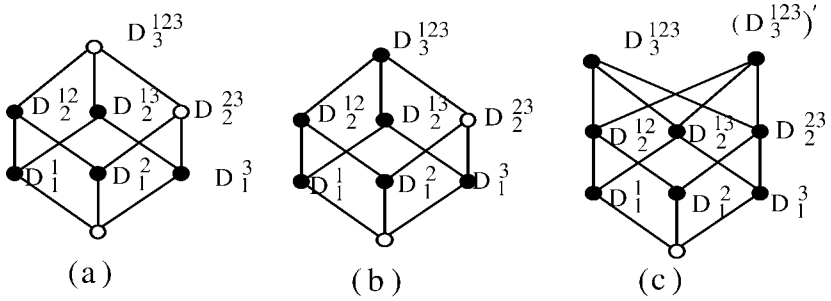

Figure 10. The VDW graphs corresponding to VDWS functions derived from figures 9(a)–(c). The vertices D_3^{123}, $(D_3^{123})'$ in figure 10(c) correspond to the vertex (1,1,1) in figure 9(c). This vertex was plit according to rule (iii).

VDW graphs for the VDWSs in figures 8(a)–(c) are given in figures 10(a)–(c) as derived according to the rules (i)–(iv).

In conclusion, it should be noted that a new model of VDWSs of molecules is proposed that differs from shape matrix code or VDW graphs, with the advantage that they are defined as finite functions whose domain is a Boolean lattice. Therefore, in contrast to the graph structure, their structural representation has a unified form. If under

VDW graphs isomorphism we understand the bijection that preserves the levels of the vertices and their incidence, then the logical models allow us to reduce the problem of VDW graph isomorphism to recognition if the corresponding logical functions belong to the same class (i.e., one can be obtained from another by permuting variables) as a consequence of the following statement.

Statement 3.3. VDW graphs are isomorphic if and only if the corresponding VDWS functions belong to the same equivalence class with respect to permutation of variables.

4. Conclusion

A new discrete mathematical model of molecular shape has been proposed that uses the partition property of a representation of molecular shape.

It was shown that the earlier discrete mathematical models such as shape matrix representations of MIDCOs or MEPCOs, as well as VDW graphs have logical analogs of the proposed type. The new model is a *unified* one, since it possesses the same form for MIDCOs, MEPCOs, as well as for VDWSs. Therefore, it is suitable for mixed quantum chemical models, particularly for interpenetrating surfaces of different natures.

The families of shapes can be logically modeled and compared by a sequence of logical models. Logical description of molecular shape structures allows the implementation of known methods of logical diagnostics to investigate the structural distinction of shapes and to introduce the metric on the set of molecular shape equivalence classes avoiding explicit embedding of objects into a vector space. These classes were defined with respect to the system of properties and symmetrical relations on the topology base that reflect the natural peculiarities of the molecular shapes.

Notice that although only logical models were introduced, the proposed approach allows us to implement many other finite functional models. For example, matrix code can be considered as a two-variable function $\{D_1, \ldots, D_n\}^2 \to \{0, 1, 2\}$. Shape families can be represented as finite functions from a space of parameter values to the set of shape descriptors. For example, (a, b)-maps of the shape groups can be considered as such functions.

Finally, we can notice, that the finite approximation of Cartesian product of the set M of internal configurations K, the set of thresholds a, and the set of curvature parameters b are partially ordered sets. Taking into account that the set of equivalence classes with respect to equivalence relation is finite as well, we can perform a mapping from that Cartesian product into the set of equivalence classes as a multivalued function. Consequently, we can implement partially ordered pattern recognition algorithms [39,40] to derive a compressed representation of these mappings and also to provide the tool for recognition of the class from which a given triplet (K, a, b) came. In the latter case the compressed representation of the mapping mentioned above is implemented as a decision rule. This possibility allows one to design compressed databases of molecular shapes to be used as decision making systems of a functional type [55,56]. These data-

bases can be created implementing the proposed logical models of molecular shapes as well as any other discrete mathematical models, mentioned in the introductory section.

Acknowledgements

Authors are grateful to Prof. A.A. Bolotov (LSI Logic, USA) for the useful consultation and discussion and to Peter Warburton (University Saskatchewan, Department of Chemistry, Canada) for linguistic help. The article was written in Collegium Budapest Institute for Advanced Study.

References

[1] P.G. Mezey, *Shape in Chemistry: An Introduction to Molecular Shape Topology* (VCH, UK, 1993).
[2] P.G. Mezey, in: *Concepts and Applications of Molecular Similarity*, eds. M.A. Jonson and G.M. Maggiora (Wiley, New York, 1990) p. 321.
[3] P.G. Mezey, in: *Reviews in Computational Chemistry*, eds. K.B. Lipkowitz and D.B. Boyd (VCH, New York, 1990) p. 265.
[4] G.A. Arteca, N.D. Grant and P.G. Mezey, J. Comput. Chem. 12 (1991) 1138.
[5] G.A. Arteca and P.G. Mezey, Theor. Chem. Acta 81 (1992) 79.
[6] C.E. Dykstra, *Ab Initio Calculation of the Structures and Properties of Molecules* (Elsevier, Amsterdam, 1988).
[7] P.G. Mezey, in: *Conceptual Trends in Quantum Chemistry*, eds. J.-L. Calis and E.S. Kryachko (Kluwer Academic, Dordrecht, 1997) p. 519.
[8] P.G. Mezey, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 51 (2000) 267.
[9] G.A. Arteca and P.G. Mezey, Internat. J. Quantum Chem. 34 (1988) 517.
[10] G.A. Arteca and P.G. Mezey, Theor. Chim. Acta 75 (1989) 333.
[11] P.G. Mezey, in: *Graph Theoretical Approaches to Chemical Reactivity*, eds. D. Bonchev and O. Mekenyan (Kluwer Academic, Dordrecht, 1994) p. 181.
[12] F. Harary and P.G. Mezey, J. Math. Chem. 2 (1988) 377.
[13] P.G. Mezey, in: *Computational Chemical Graph Theory*, ed. H. Rouvray (Nova, New York, 1990) p. 175.
[14] A.T. Balaban, in: *Chemical Topology to Three Dimensional Geometry*, ed. A.T. Balaban (Plenum, New York, 1997) p. 1.
[15] L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research* (Academic Press, New York, 1976).
[16] M. Randic, Acta Chim. Slov. 44 (1997) 57.
[17] D. Bonchev, J. Mol. Struct. 336 (1995) 137.
[18] P.G. Mezey, Internat. J. Quantum Chem. Quant. Biol. Sympos. 14 (1987) 127.
[19] P.G. Mezey, J. Math. Chem. 2 (1988) 299.
[20] G.A. Arteca and P.G. Mezey, J. Math. Chem. 3 (1989) 43.
[21] A. Badel-Chagnon, P. Tuffery, S. Hazout and P.G. Mezey, Molec. Engrg. 8 (1999) 267.
[22] F. Harary and P.G. Mezey, Theor. Chim. Acta 79 (1991) 379.
[23] F. Harary and P.G. Mezey, Discrete Appl. Math. 19 (1988) 205.
[24] P.G. Mezey, J. Math. Chem. 11 (1992) 27.
[25] P.G. Mezey, Bolyai Soc. Math. Stud. 7 (1999) 323.
[26] P.G. Mezey, in: *From Chemical Topology to Three-Dimensional Geometry*, ed. A.T. Balaban (Plenum, New York, 1997) p. 25.
[27] P.G. Mezey, Internat. J. Quantum Chem. 51 (1994) 255.

[28] F. Harary and P.G. Mezey, Internat. J. Quantum Chem. 62 (1997) 353.
[29] I.A. Chegis and S.V. Yablonsky, Trudy Mat. Inst. Steklov. 51 (1958) (in Russian).
[30] E.J. McCluskey, *Logic Design Principles* (Prentice-Hall, Englewood Cliffs, NJ, 1986).
[31] A.B. Frolov, *Models and Methods of Technical Diagnostics* (Znanie, Moscow, 1990) (in Russian).
[32] L. Goldfarb and A.K.C. Wong, in: *Proceedings of 17th IEEE Decision and Control Conference* (1979) p. 1086.
[33] A.K.C. Wong and L. Goldfarb, in: *Pattern Recognition Theory and Applications*, eds. J. Kittler and L.F. Pau (Reidel, Dordrecht, 1982) p. 157.
[34] E. Jako, Development and analysis of methods for canonical logical partitioning of Boolean functions: Applications in systems of logical design, PhD Thesis, MPEI, Moscow (1983).
[35] E. Jako and I. Fabian, Hungarian patent 203173, INT CL:H03K 19/173 (1989).
[36] A.B. Frolov and E. Jako, Trudy Moskov. Energ. Inst. 555 (1981) 125 (in Russian).
[37] E.A. Yakubajtis, Avtomatika i Vitshislitelnaja Technika 1 (1974) 1 (in Russian).
[38] I.I. Zhegalkin, Mat. Sb. 34(1) (1927) 9 (in Russian).
[39] A.B. Frolov and E. Jako, Izvestiya Akademii Nauk SSSR. Tekhnicheskaya Kibernetika 5 (1990) 96 (in Russian).
[40] A. Frolov and E. Jako, *Algorithms for Recognition of Partially Ordered Objects and Their Application* (Scripta Technica, Wiley, 1991).
[41] E. Jako and P. Ittzés, Abstracta Botanica 22 (1998) 121.
[42] E. Jako, Ordered sets in molecular biology, in: *Proc. of Beckman Res. Inst., XVIth Annual Advance*, California (1996).
[43] A.A. Bolotov and A. Shajeb, DAN SSSR (1987) 297; 3 (1987) 527 (in Russian).
[44] A.A. Bolotov, Diskretnaya Matematika 8(4) (1996) 62 (in Russian).
[45] A.A. Bolotov and A.B. Frolov, *Classification and Recognition in Discrete Systems* (MPEI, Moscow, 1997).
[46] P.G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987).
[47] G.A. Arteca and P.G. Mezey, J. Comput. Chem. 9 (1988) 554.
[48] G.A. Arteca, V.B. Jammal, P.G. Mezey, J.S. Yadav, M.A. Hermsmeier and T.M. Gund, J. Mol. Graphics 6 (1988) 45.
[49] S.V. Yablonsky, *Introduction to Discrete Mathematics* (Mir, Moscow, 1988).
[50] R. Lidl and G. Pilz, *Applied Abstract Algebra* (Springer, Berlin, 1984).
[51] K.H. Rosen, *Discrete Mathematics and Its Applications* (McGraw-Hill, New York, 1991).
[52] A. Mishchenko and A. Fomenko, *A Course of Differential Geometry and Topology* (Mir Publisher, Moscow, 1988).
[53] A.V. Aho, J.E. Hopcroft and J.D. Ulman, *The Design and Analysis of Computer Algorithms* (Addison-Wesley, Reading, MA, 1976).
[54] A.E. Andreev, A.A. Bolotov and A.B. Frolov, *Discrete Mathematics: Discrete Optimisation Problems and Complexity of Algorithms* (MPEI, Moscow, 2000).
[55] A.B. Frolov and D.A. Frolov, Vestnik MEI 6 (1996) 113 (in Russian).
[56] A.B. Frolov, D.A. Frolov and I.D. Chetrafilov, Vestnik MEI 2 (1996) 201 (in Russian).