Foundations of Science (2022) 27:1327–1345
https://doi.org/10.1007/s10699-022-09830-8

The Information Encoded in Structures: Theory and Application to Molecular Cybernetics

Andrzej Bielecki[1] · **Michael Schmittel**[2]

Accepted: 14 January 2022 / Published online: 16 February 2022

Abstract
Theoretical frames for analyzing information in biological and molecular multicomponent structures are proposed. The mathematical foundations of the proposal are presented. Both the information encoded in structures is defined and the method of calculating the amount of this information is introduced. The proposed approach is applied to the operation of a molecular multicomponent machine.

Keywords Structure · Relation · Structural information · Molecular machines · Molecular cybernetics

1 Introduction

Information, together with matter and energy, is contemporarily regarded as a fundamental component of the existing world—see, for instance, Barreiro et al. (2020), Krzanowski (2020a) and references given there. This implies a widely postulated thesis that specific information processing, in various contexts and on various levels, is an immanent property of living beings. This processing takes place not only on the level of the nervous system and a single neuron—see, for instance, (Bielecki, 2019; Sadeh & Clopath, 2020; Schepherd, 1994; Tadeusiewicz, 2010; Rosslenbroich, 2014), chapter 8, but also on the levels of biological systems and subcellular organelles and processes (Bielecki, 2015; Hellerman, 2016; Jablonka & Lamb, 2006; Nurse, 2008; Perez Velazquez, 2009), including viruses (Rohwer & Barott, 2013). The problem is that, so far, there is no adequate information theory that can be used in biology, although the problem is considered—see, for instance (Smith, 2000; Walker et al., 2016). First of all, the existing information theories—Shannon and Kolmogorov theories can be put as the best-known examples—provide the measure for the amount of information rether

✉ Andrzej Bielecki
bielecki@agh.edu.pl; azbielecki@gmail.com

Michael Schmittel
schmittel@chemie.uni-siegen.de

[1] Faculty of Electrical Engineering, Automation, Computer Science and Biomedical Engineering, AGH University of Science and Technology, Mickiewicza 30, 30-059 Cracow, Poland

[2] Center of Micro and Nanochemistry and Engineering, Organische Chemie I, Universität Siegen, Adolf-Reichwein-Str. 2, 57068 Siegen, Germany

than tools for analyzing of information type and specifics, except the probabilistic aspect of information in the case of Shannon theory. Shanon information theory adequately describes the transmission of signals over a channel and, therefore, it can be used successfully to analyze information in digital-type chains, such as genetic code (Schneider, 2000). It is useless, however, for the analysis of information encoded in non-sequential structures. This is a key disadvantage because this type of information, being a special type of ontological information, is crucial in biology (Krzanowski, 2020b).

The Hellerman proposal (Hellerman, 2006) that is dedicated to the analysis of the level of organization in structures as such, first of all, biological ones, seems to be a proper starting point for working out the adequate biological information theory. The fact that information in organisms is encoded by biological structures—see Bielecki (2015), Jablonka and Lamb (2006) for details—is the most important reason for the adequacy of Hellerman approach. In the current state of development of the presented theory, the simplest possible level should be used to test the efficiency of the proposed approach. The supramolecular cybernetics, including molecular machines controlled by molecular switches (Biswas et al., 2020; Goswami et al., 2019; Paul et al., 2019; Schmittel, 2019), is the most appropriate level for testing the theory at its current level of development. On the one hand, there are a lot of examples of molecular machines that are relatively simple structures, at least in comparison with most biological structures. Such machines are suitable for testing the proposed theory at its current, initial stage. On the other hand, dynamic biological structures are molecular machines—the ribosome can be put as an example (Rodnina & Wintenmeyer, 2011; Spirin, 2002, 2009).

This publication presents a theoretical framework for defining and analyzing information encoded in the structure of a multicomponent ensemble and it is a far-reaching extension of the context in which the problem of coding information with the use of physical structures is normally considered—see, for instance, (Goldfeld et al., 2021). It should be emphasized that the existence of this type of information results only from the fact that a given structure has such and no other form. The presented approach is a development of the ideas sketched in Bielecki (2015), sections 5.2 and 5.3, and is the realization of the first stage of the research program outlined there including *introduction of formal definitions of the proposed terms and studying their properties and relations between them. This is related to creating the formalization which describes biological phenomena properly.* It should be stressed that lack of such formalization was emphasized in Perez Velazquez (2005). The introduced formalization was applied to calculate amount of information contained in the molecular machines described in Paul et al. (2019).

The paper is organized in the following way. The proposed mathematical formalization is presented in Sect. 2. It should be emphasized that the proposal includes only the formalization of static, non-hierarchical structures. The theory of dynamic structures, as well as hierarchical ones, is planned to be the topic of the subsequent papers. Foundations of molecular cybernetics are briefly presented in Sect. 3, whereas application of the proposed theory to molecular machines is put forward in Sect. 4. Potential utility of the propsed approach is discussed in Sect. 5.

2 Formulation of Information Theory

In this section formalization of the concept of information contained in a structure is proposed. Furthermore, the way the amount of such information can be measured is presented. The approach is formulated in a pure mathematical aspect. The authors by no means reject

the concept of information as an entity. At the current stage of the studies, however, the primary goal was to develop a mathematical basis of structural information. The enhancement of the concept with physical aspects will be the next step in the proposed approach and it is planned to be a topic of subsequent papers.

In this paper, only information on finite sets is studied. The very idea can be put forward as follows:

(a) There is no information in general. Information always exists in a given set.
(b) Information is generated by relations defined on this set.
(c) The generated information is unambiguously related to a certain graph.
(d) Both the information as such and the amount of information are determined by the structure of this graph.

Let us start with recalling the classical way the mentioned graph is generated—see, for instance, (Carnap, 1928), paragraph 11.

Definition 2.1 Let the set X consist of the elements $x_1, \ldots, x_n$ and let $\mathcal{R}$ be a relation on X. The oriented graph (orgraph) $G := \{V, E\}$, where V is the set of its nodes and E is the set of its oriented edges, is the graph generated by the relation $\mathcal{R}$ if $V = X$ and $(x_i, x_j) \in E$ iff $x_i \mathcal{R} x_j, i, j \in \{1, \ldots, n\}$.

The graph generated by the relation $\mathcal{R}$ on the set X will be denoted, in the sequel, by $G(X, \mathcal{R})$.

Definition 2.1 simply implies the following corollary.

Corollary 2.2 *A finite set X and a relation $\mathcal{R}$, defined on X, generate a unique oriented graph (orgraph). Furthermore, each orgraph $G := \{V, E\}$ generates a unique relation $\mathcal{R}$ on the set $X = V$ in the following way: if $(x_i, v_j) \in E$, then $x_i \mathcal{R} x_j$.*

In the sequel, it is assumed that $\mathcal{R}$ is an anti-reflexive relation on X i.e. $\forall_{x \in X} \neg(x \mathcal{R} x)$. This means that the graph G has no loops, i.e. the edges of the form (x, x) because no element is in relation with itself. A set X with a relation $\mathcal{R}$ will be denoted, in the sequel, as $(X, \mathcal{R})$.

Let us define node-balls on orgraphs.

Definition 2.3 Let $G = (V, E)$ be an orgraph. A node-ball $B_G^{node}(x, 1)$ of radius 1 and a center at the node $x \in V$ is a subgraph (V_1, E_1) such that $V_1 := \{y \in V : y = x \vee (x, y) \in E \vee (y, x) \in E\}$ and $E_1 := \{(u, v) \in E : u \in V_1 \wedge v \in V_1\}$. A ball $B_G^{node}(x, n)$ of radius $n \in \{2, 3, \ldots\}$ is a union

$$B_G^{node}(x, n) := \bigcup_{y \in X} B_G^{node}(y, 1),$$

where $y \in B_G^{node}(x, n-1)$. Let us also put $B_G^{node}(x, 0) := (\{x\}, \emptyset)$. Let us also define $r_{node}(G, x)$ as a minimal natural number such that $B_G^{node}(x, r_{node}(G, x)) = B_G^{node}(x, r_{node}(G, x) + 1)$.

Intuitively, $r_{node}(G, x)$ is a maximal radius of the node-ball with the center at the node x which means that for this radius the ball is equal to the graph G connected component that contains the node x.

Let us define node information.

Definition 2.4 Node information I^{node} on X introduced by $\mathcal{R}$ is a set of all node-balls on $G(X, \mathcal{R})$, i.e. $I^{node} := \left\{B_G^{node}(x, n),\ x \in V,\ n \in \{0, \ldots, r_{node}(G, x)\}\right\}$.

The node information is the set of all node-balls in the graph because based on this set the elements of the set X are identified. This identification is carried out by distinguishing the nodes of the graph that represent the elements of the set X.

Let us define indistinguishable nodes of a graph $G(X, \mathcal{R})$.

Definition 2.5 Let $G = (V, E)$ be an orgraph and let $x, y \in V$. The nodes x and y are indistinguishable if they belong to the same connected component of G and for each $n \in \mathbb{N}^+$ the balls $B_G^{node}(x, n)$ and $B_G^{node}(y, n)$ are isomorphic. Otherwise the nodes are distinguishable.

Let us define indistinguishability by $\mathcal{R}$ of two elements $x, y \in X$ in $(X, \mathcal{R})$.

Definition 2.6 Let $(X, \mathcal{R})$ be given. The elements $x, y \in X$ are indistinguishable by $\mathcal{R}$ if they are indistinguishable as the nodes of $G(X, \mathcal{R})$.

Let us define the indistinguishability relation $\mathcal{D}^{node}(X, \mathcal{R})$ for a given $(X, \mathcal{R})$.

Definition 2.7 For $x, y \in X$, $x\mathcal{D}^{node}(X, \mathcal{R})y$ if the element x is indistinguishable from the element y by $\mathcal{R}$.

It is obvious that the relation $\mathcal{D}^{node}$ is an equivalence relation on X. The equivalence classes of $\mathcal{D}^{node}$ define the organization of the set X in the Hellermann sense (Hellerman 2006). Thus, the amount of node information generated by $\mathcal{R}$ on X is given by the formula introduced by Hellerman:

$$H^{node} = -n \sum_{k=1}^{K} \frac{n_k}{n} \log \frac{n_k}{n}, \tag{1}$$

provided that n is a number of elements in the set X, K is a number of elements of the quotient set $X/\mathcal{D}^{node}$, and n_k is the number of elements in the $k-$th equivalence class. This means that the said organization of the set X in the Hellermann sense has the form $(n_1, \ldots, n_K)$, where $n_1 + \cdots + n_K = n$.

In formula (1) and in the sequel log denotes $\log_2$.

Let us introduce the labeling of the nodes. Let L be an equivalence relation on X.

Definition 2.8 The canonical projection $f_L(x) = [x]_L$ is the labeling function of the set X.

Remark 2.9 Classically labeling is done by using the labeling function. In this paper, it is done by using an equivalence relation and canonical projection so that the labeling is not arbitrary, but is implied by the properties of the elements of the set X. For example, in a chemical molecule, the atoms of the same chemical element will have the same label.

In the sequel, labeling be the only relation that will be denoted by node labels, not by the edges of the graph—see Sect. 4.

Let us define a structure.

Definition 2.10 Let $X = \{x_1, \ldots, x_n\}$ be a finite set and let $\mathcal{R}_1, \ldots, \mathcal{R}_n$ be relations on X. The $n+1$-tuple S :

$$S(X, n) := (X, \mathcal{R}_1, \ldots, \mathcal{R}_n),$$

is said to be a structure on X.

Each relation generates an orgraph on X in the way described in Definition 2.1. In order to define amount of node information in the structure $S := (X, \mathcal{R}_1, \ldots, \mathcal{R}_n)$ let us recall a simple property.

Corollary 2.11 *Let $\mathfrak{A} = \{\mathcal{A}_t\}_{t\in T}$ and $\mathfrak{B} = \{\mathcal{B}_s\}_{s\in S}$ be partitions of the set X. Let us define $\mathcal{C}_{t,s} := \mathcal{A}_t \cap \mathcal{B}_s$, $t \in T, s \in S$. Then, the family $\mathfrak{C} := \{\mathcal{C}_{t,s} : \mathcal{C}_{t,s} \neq \emptyset\}$ is a partition of X.*

The above property remains true for a finite number of partitions.

In the structure $S(X, n)$ the equivalence classes $\mathcal{D}_i$, $i \in \{1, \ldots, n\}$, created by relations $\mathcal{R}_1, \ldots, \mathcal{R}_n$ generate n partitions of X. The partitions are constituted by the equivalence classes of $\mathcal{D}_i$—see Definition 2.7 Let $\mathfrak{Y}$ be the family of all nonempty sets of the form

$$\mathfrak{Y} := \left\{Y_{k_1,\ldots,k_n} := Y_{k_1} \cap \cdots \cap Y_{k_n}\right\},$$

where Y_{k_i} is the k_i-th equivalence class generated by $\mathcal{D}_i$. By Corollary 2.11 the family $\mathfrak{Y}$ is a partition of X.

Definition 2.12 The amount of node information generated by $\mathcal{R}_1, \ldots, \mathcal{R}_n$ on $S(X, n)$ is given by the formula:

$$H^{node} = -n \sum_{k=1}^{K} \frac{n_k}{n} \log \frac{n_k}{n}, \tag{2}$$

provided that n is the number of elements in the set X, K is a number of elements of the family $\mathfrak{Y}$ and n_k is the number of elements in the $k-$th set of $\mathfrak{Y}$.

Formula (2) tells how pieces of information generated by various relations overlap.

Let us notice that node information is insufficient to describe information that is represented by a graph generated by a relation. It is obvious that the amount of information in a cyclic graph is different from the amount of information in a fully-connected graph that has the same number of nodes. Therefore it is necessary to introduce information represented by edges.

Edge information can be defined in a similar way as node information. Thus, let us define edge-balls on orgraphs.

Definition 2.13 Let $G = (V, E)$ be an orgraph. An edge-ball $B_G^{edge}(e, 1)$ of radius 1 and a center at the edge $e = (u, v) \in E$ is a subgraph (V_1, E_1) such that $V_1 := \{u, v\}$ and $E_1 := \{(u, v), (v, u)\} \cap E$. A ball $B_{ed}^G(e, n)$ of radius $n \in \{2, 3, \ldots\}$ is a union

$$B_G^{edge}(e,n) := \bigcup_{c\in E} B_G^{edge}(c,1),$$

where $(c=(x,y)\vee c=(y,x))\wedge y\in V_{n-1}$, provided that $B_G^{edge}(e,n-1)=(V_{n-1},E_{n-1})$. Let us also define $r_{edge}(G,e)$ as a minimal natural number such that $B_G^{edge}(e,r_{edge}(G,e))=B_G^{edge}(e,r_{edge}(G,e)+1)$.

Let us define edge information.

Definition 2.14 Edge information I^{edge} on X introduced by $\mathcal{R}$ is a set of all edge-balls on $G(X,\mathcal{R})$, i.e. $I^{edge} := \left\{ B_G^{edge}(e,n),\ e\in E,\ n\in\{0,\dots,r_{edge}(G,e)\}\right\}$.

The edge information is the set of all edge-balls in the graph because this set provides information about the variety of relations on X.

Let us define indistinguishable edges of a graph $G(X,\mathcal{R})$.

Definition 2.15 Let $G=(V,E)$ be an orgraph and let $e_1,e_2\in E$. The edges e_1 and e_2 are indistinguishable if they belong to the same connected component of G and for each $n\in\mathbb{N}^+$ the balls $B_G^{edge}(e_1,n)$ and $B_G^{edge}(e_2,n)$ are isomorphic. Otherwise the edges are distinguishable.

Let us define indistinguishability by $\mathcal{R}$ of two elements $e_1,e_2\in E$ in $G(X,\mathcal{R})$.

Definition 2.16 Let $G(X,\mathcal{R})$ be given. The elements $e_1,e_2\in E$ are indistinguishable by $\mathcal{R}$ if they are indistinguishable as the edges of $G(X,\mathcal{R})$.

Let us define the edge indistinguishability relation $\mathcal{D}^{edge}(X,\mathcal{R})$ for a given $G(X,\mathcal{R})$.

Definition 2.17 For $e_1,e_2\in G(X,\mathcal{R})$ $x\mathcal{D}^{edge}(X,\mathcal{R})y$ if the element e_1 is indistinguishable from the element e_2.

Relation $\mathcal{D}^{edge}$ is an equivalence relation on E. The equivalence classes of $\mathcal{D}^{edge}$ define the organization of the set E in the Hellermann sense (Hellerman 2006). In the case of edges, the amount of information should be calibrated by all possible edges, not only by the total number of the existing ones. Thus, let N be the number of all possible edges, $N=n(n-1)$, where n is a number of elements in X. Let us remember that $\mathcal{R}$ is antireflexive which means that there are no loops. Therefore n is multiplied by $n-1$, not by n. The amount of edge information generated by $\mathcal{R}$ on E is given by the modified Hellerman formula:

$$H^{edge} = -n\sum_{k=1}^{K}\frac{n_k}{N}\log\frac{n_k}{N}, \tag{3}$$

provided that K is a number of elements of the quotient set $E/\mathcal{D}^{edge}$, and n_k is the number of elements in the $k-$th equivalence class. Let us denote the modified Hellerman organization as $(n_1,\dots,n_K)(N)$, where $n_1+\cdots+n_K=n$.

The total information contained in the structure is given by the formula

$$H = H^{node} + H^{edge} \tag{4}$$

To sum up, the labeling information tells which elements make up the structure—see Definition 2.8 and Remark 2.9. The node information generated by the remaining relationships, on the other hand, tells how the elements relate to each other. Edge information tells how many nodes are in a relation and what is the specifics of relationships between the elements.

3 Supramolecular Cybernetics: Molecular Machines and Switches

Although artificial molecular machines have received ample attention over the past 30 years and despite the spotlight that the awarding of the Nobel Prize has inevitably given this field (Sauvage, Stoddart, and Feringa in 2016) (Boyer, 1998), they are predominantly designed and studied as stand-alone devices. Cybernetic aspects, which play a vital role in the operation of natural machines, have so far been completely left out (Schmittel & Howlader, 2020). This deficit is rooted in the extreme complexity of even small cybernetic circuitry that has so far eluded any planning approach. It has thus been emphasized that any developments in this area will depend on the systematic expansion and improvement of molecular communication, which requires a broad range of purposeful molecular receptors, switches, and regulators. Only with those molecular constituents in hand one may be able to build cybernetic control loops and feedback mechanisms in a bottom-up approach.

3.1 Artificial Molecular Machines

Let us start with a thought experiment on a simple molecular machine. The minimum requirements for its cybernetic control involve control of the directionality of the movement (for instance, unidirectional or forward vs. backward) and the provision of energy for the regulation of the motional amplitude. There are spectacular machines alternately driven by light and thermal energy (van Leeuwen et al., 2017) as well as pumps (Qiu et al., 2020) in turns operated by redox potential changes and thermal energy (Baroncini et al., 2020). In those cases, light/redox processes allow overcoming a high barrier whereas the thermal energy enables a directional motion by overcoming a small barrier. In contrast, the cybernetic regulation of biological machines is achieved by a gamut of networked and interdependent chemical inputs including a chemical form of energy (ATP), all of which copperate in adapting the machine's operation to the environmental needs. While recently molecular machines have successfully been driven with chemical energy, it would be important to influence their modus operandi with further control variables. The Schmittel group has focused on wiring molecular machines to molecular switches that influence the state of the machine, actually in many cases catalytic machinery, with the goal to later drive them with chemical energy (Schmittel & Howlader, 2020). At present, the machines operate on thermal energy whereas directionality is controlled through molecular switching via chemical signaling. In some instances, a level of information handling was reached that characterizes smart mixtures (Goswami et al., 2019). Without going into detail, an ensemble of 13 constituents was designed to respond to the availability of a single outside trigger in a way that two different catalytic processes were initiated, one in the presence of the trigger, another in its absence.

3.2 The Studied Case

In the next section, the theoretical proposal introduced in Section 2 is applied to analyze information in the molecular machine described in Paul et al. (2019). Let us briefly recall the results: In this work, a networked seven-component ensemble is utilized to control initially the synthesis and afterward the shuttling of a [2]rotaxane. Operation of the networked system was achieved by the addition and removal of zinc(**II**) ions that internally triggered copper release the latter being monitored by distinct colorimetric and fluorimetric signals. The mode of the operation depended decisively on communication within the ensemble where the chemical signal, i.e. copper(**I**) ion, exerted two different roles:

(a) In the beginning it served as a catalyst to build a [2]-rotaxane (NetState **I** → **II**),
(b) Later it enabled switching ON/OFF of the thermal oscillations in the rotaxane (NetState **II** ⇆ **III**).

Addition of zinc ions led to liberation of copper(**I**) ions from the dimeric release system (Fig. 1). The dimeric aggregate had been chosen, because the assemblies $[Cu_2(\mathbf{1})_2]^{2+}$, $[Zn_2(\mathbf{1})_2]^{4+}$ and ligand **1** displayed different emissions in the fluorescence channel allowing to monitor changes in the overall system.

Upon the first addition of zinc(**II**) ions (2 equiv.) the release system $[Cu_2(\mathbf{1})_2]^{2+}$ discharged two equiv. of copper(**I**) ions that acted as catalyst to afford two equiv. of the copper-loaded [2]-rotaxane $[Cu(\mathbf{5})]^+$. Due to the thermal energy at room temperature the macrocycle in $[Cu(\mathbf{5})]^+$ was found to oscillate at $k_{298} = 30$ kHz between the two triazole stations (Fig. 2). At this point, removal of zinc(**II**) freed $[Cu(\mathbf{5})]^+$ from copper(**I**) ions and regenerated $[Cu_2(\mathbf{1})_2]^{2+}$ The copper-free [2]-rotaxane **5** did not show any oscillations (on the NMR time scale).

The complete picture in Fig. 3 describes how the three networked states (NetStates) arise from combing all constituents in the proper stoichiometry prior to addition and after removal of zinc(**II**) ions. An important issue is that the initial addition of zinc(**II**) triggers an irreversible formation of the [2]-rotaxane (NetState **I→II**), whereas alternate addition and removal of zinc then allows shuttling the system between NetStates II and III. Removal of zinc(**II**) was readily accomplished by the addition of hexacyclene.

Fig. 1 Liberation of copper(**I**) ions from the dimeric $[Cu_2(\mathbf{1})_2]^{2+}$ upon addition of zinc(**II**) ions

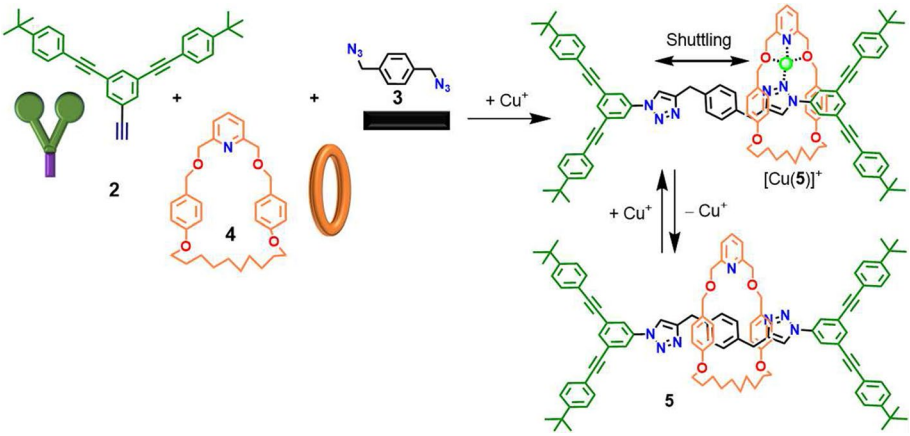


Fig. 2 The copper(**I**)-catalyzed formation of the copper(**I**)-loaded rotaxane and the reversible copper(**I**) removal and addition

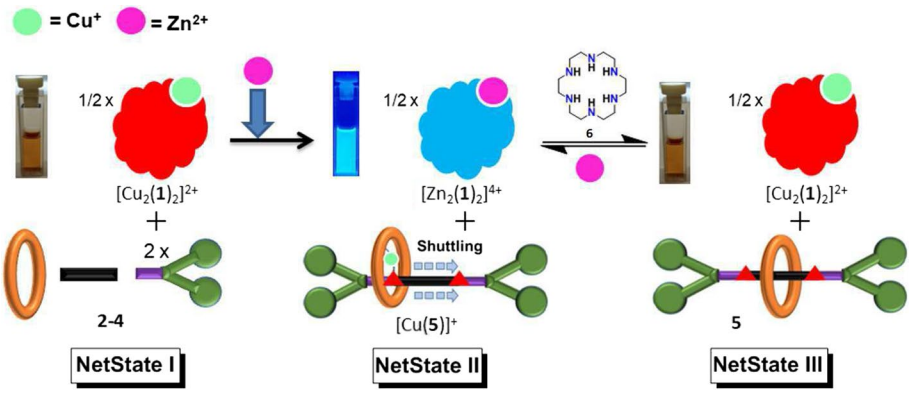


Fig. 3 The full cybernetic system connecting three networked states

4 Information in Supramolecular Cybernetics: Static Case

The proposed approach to information encoded in structures has been applied to the molecular machine, described in Paul et al. (2019) and briefly recalled in Sect. 3.2. It should be stressed, that although in the previous section as well as in Paul et al. (2019) the dynamic aspect is discussed as well—the full cybernetic system dynamics (Figs. 1, 2, 3) and shuttling of [Cu (**1**)(**5**)]$^{+}$—in this paper only the static aspect is considered. This means, that we study only information encoded in structures. The problem of information processing during synthesis, self-sorting, and structural transformations, including stoichiometric aspects, will be studied in the next paper.

It should be also stressed that only information of the highest level of organization is taken into account in our calculations below. This means that only relations between substructures that form the final structures are considered i.e. information contained in the said substructures is not included. Moreover, let us ephasize that in below

considerations, we use a large simplification by introducing only one relation: *"element α is connected to the element β"* without taking into account various types of connections between elements.

The presented application covers the case where the nodes of the generated graph are labeled and, apart from the labeling relation, only one additional relation on the set is established. Thus, the set $X = \{x_1, x_2, x_3, \ldots x_{12}\}$ consists of twelve elements that are chemical compounds. Let us introduce the labeling in such a way, that labels correspond to chemical compounds, according to Definition 2.8 and Remark 2.9. Thus, let the labeling relation be denoted as L and let $x_i L x_j$ if x_i and x_j are the same chemical compounds. After labeling the set X has the following form:

$$\{\text{ tr, tr, arm, arm, hex, hex, Cu, Cu, Zn, Zn, rg, bar }\}.$$

The order of the specified above labels corresponds to the order of the indices of the elements which means that x_1 has the label **tr**, i.e. lab $(x_1) =$ tr , lab $(x_2) =$ tr , lab $(x_3) =$ arm , lab $(x_4) =$ arm , lab $(x_5) =$ hex , lab $(x_6) =$ hex , lab $(x_7) =$ Cu , lab $(x_8) =$ Cu , lab $(x_9) =$ Zn , lab $(x_{10}) =$ Zn , lab $(x_{11}) =$ rg , lab $(x_{12}) =$ bar . The used labels correspond to the following chemical compounds:

tr	denotes ligand **2**—see Fig. 2 and Paul et al. (2019)
arm	denotes ligand **4**—see Fig. 2 and Paul et al. (2019)
hex	denotes hexacyclen,
Cu	denotes $[\,Cu(CH_3\,CN)_4)PF_{\,6}]$,
Zn	denotes $Zn(OTf)_2$,
rg	denotes the macrocycle—see Fig. 2 and Paul et al. (2019),
bar	denotes the phenyl ring as part of diazide **3**—see Fig. 2 and Paul et al. (2019).

The graphical symbols of the above structures, the same that have been used in Paul et al. (2019), are shown in Fig. 4. As it has been mentioned at the beginning of this section, the inner structure of the above compounds is not taken into consideration. Thus, the labeling relation L generates the following partition of X : (2, 2, 2, 2, 2, 1, 1) and, as a consequence, it creates the following amount of information:

$$H = -12\left[5\left(\frac{2}{12}\log\frac{2}{12}\right) + 2\left(\frac{1}{12}\log\frac{1}{12}\right)\right] = 33.007.$$

Apart from the defined labeling relation L only one relation, let us say R, is defined on X : $x_i R x_j$ if x_i is connected to x_j.

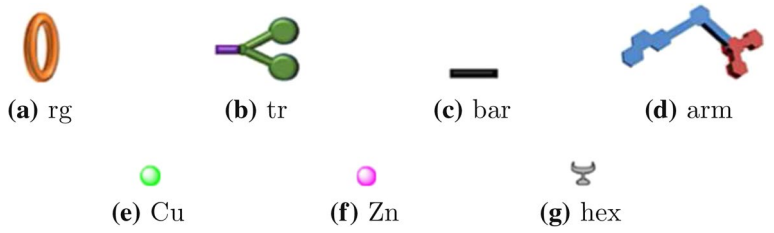
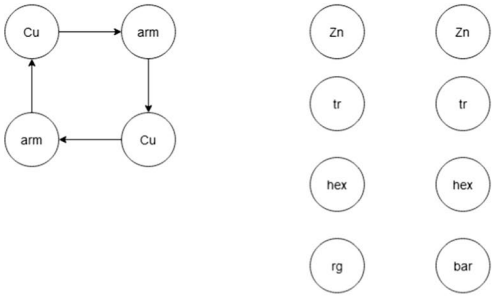


Fig. 4 Graphical symbols corresponding to the chemical compounds

(a) the created structure



(b) relations

Fig. 5 NetState I (NS-I)

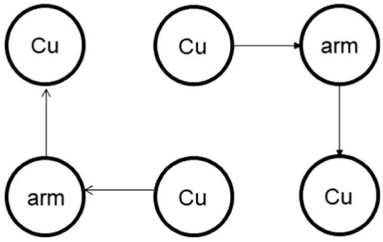


Fig. 6 Two isomorphic node-balls $B_G^{node}(\mathbf{arm}, 1)$ in NS-I

Let us consider the NetState I (NS-I, for abbreviation)—see Fig. 5a. The following elements are in relations:

$$x_7 R x_3,\ x_8 R x_4,\ x_3 R x_8,\ \text{ and } x_4 R x_7,$$

which generates the graph $G_{NS-I}(X, L, \mathcal{R})$ shown in Fig. 5b. For both nodes labeled as **arm** all node-balls are isomorphic (see Fig. 6) becuse in both cases

$$B_G^{node}(\mathbf{arm}, 1) = (\{\mathbf{arm}, \mathbf{Cu}, \mathbf{Cu}\}, \{(\mathbf{Cu}, \mathbf{arm}), (\mathbf{arm}, \mathbf{Cu})\}),$$

see Fig. 6. Similarly, in both cases, $B_G^{node}(\mathbf{arm}, 2)$ is equal to the whole connected component of the graph containing the nodes labeled with labels **arm** and **Cu**. We have a similar situation in the case of node-balls with the center in the nodes labeled as **Cu**. Thus, for both nodes labeled by **Cu** and **arm** all balls are isomorphic (see Definitions 2.3 and 2.5), the relation $\mathcal{R}$ does not enable to distinguish nodes in both mentioned pairs. Therefore the relation $\mathcal{R}$ does not introduce the possibility of distinguishing additional elements of X in relation to the distinction made by labeling and, as a consequence, in the graph $G_{NS-I}(X, L, \mathcal{R})$ no two nodes that have the same label can be distinguished. As a result, we get

$$H^{nodes}(\text{ NS-I}) = 33.007.$$

In $G_{NS-I}(X, L, \mathcal{R})$ there are four edges—two (arm , Cu) and two (Cu , arm). Both edges (arm , Cu) are indistinguishable. Indeed, in both cases (see Fig. 7) we have

$$B_G^{edge}((\mathbf{arm}, \mathbf{Cu}), 1) = (\{\mathbf{arm}, \mathbf{Cu}\}, \{(\mathbf{arm}, \mathbf{Cu})\}),$$
$$B_G^{edge}((\mathbf{arm}, \mathbf{Cu}), 2) = (\{\mathbf{Cu}, \mathbf{arm}, \mathbf{Cu}\}, \{(\mathbf{Cu}, \mathbf{arm}), (\mathbf{arm}, \mathbf{Cu}), (\mathbf{Cu}, \mathbf{arm})\})$$

and, in both cases, edge-ball $B_G^{edge}((\mathbf{arm}, \mathbf{Cu}), 3)$ is equal to the whole connected component of the graph containing the nodes labeled with labels **arm** and **Cu**. We have a similar situation in the case of edge-balls with the center in the edges (**Cu, arm**). The number of all possible edges is equal to $12 \cdot 11 = 132$ (let us remember that R is anti-reflexive which means that there are no loops; therefore 12 is multiplied by 11, not by 12.) Thus, the pairs that are in relation generate the partition of the set of edges: (2, 2)(132), so

$$H^{edges}(\text{ NS-I}) = -4 \cdot \left[2 \cdot \left(\frac{2}{132} \cdot \log \frac{2}{132}\right)\right] = 0.725$$

and—see formula (4)

$$H(\text{ NS-I}) = 33.732.$$

Let us consider NS-II—see Fig. 8a. The following elements are in relations:

$$x_9 R x_3,\ x_{10} R x_4,\ x_3 R x_{10},\ x_4 R x_9,\ x_7 R x_{11},\ x_{11} R x_{12},\ x_{12} R x_1 \text{ and } x_{12} R x_2,$$

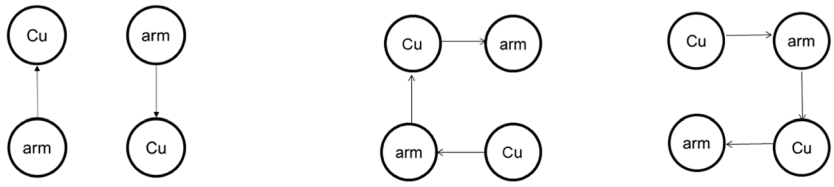
which generates the graph $G_{NS-II}(X, L, R)$ shown in Fig. 8b. The specified relation allows us to distinguish **Cu** elements because one is free whereas **rg** element is connected to the second one. Thus, in relation to the distinction made by labeling, the additional distinction in the set of the graph nodes is created and the partition of the set of nodes has the following form: (2, 2, 2, 2, 1, 1, 1, 1). As a consequence

$$H^{nodes}(\text{ NS-II}) = -12 \cdot \left[4 \cdot \left(\frac{2}{12} \cdot \log \frac{2}{12}\right) + 4 \cdot \left(\frac{1}{12} \cdot \log \frac{1}{12}\right)\right] = 35.000.$$

In $G_{NS-II}(X, L, R)$ there are three pairs of indistinguishable edges: (arm , Zn), (Zn , arm), and (bar , tr). Furthermore, there are two single distinguishable edges—(rg , Cu) and (Cu , bar). This creates in the set of edges the partition (2, 2, 2, 1, 1)(132), so

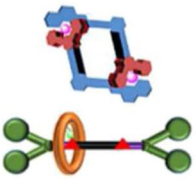
$$H^{edges}(\text{ NS-II}) = -8 \cdot \left[3 \cdot \left(\frac{2}{132} \cdot \log \frac{2}{132}\right) + 2 \cdot \left(\frac{1}{132} \cdot \log \frac{1}{132}\right)\right] = 3.030$$

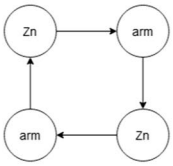
and

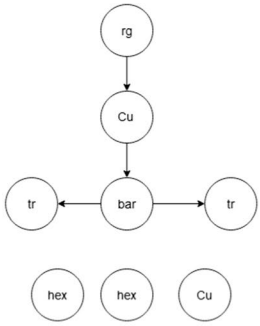


(a) Two edge-balls $B_G^{edge}((\mathbf{arm}, \mathbf{Cu}), 1)$ **(b)** Two edge-balls $B_G^{edge}((\mathbf{arm}, \mathbf{Cu}), 2)$

Fig. 7 Edge-balls in NS-I. The balls presented in **a** are isomorphic as well as the balls in **b**







(a) the created structures **(b)** relations

Fig. 8 NetState II (NS-II)

$$H(\text{ NS-II}) = 38.030.$$

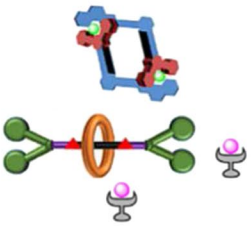
Let us consider NS-III—see Fig. 9a. The following elements are in relations:

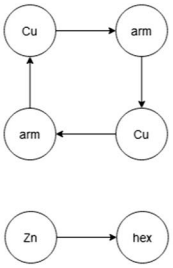
$$x_7 R x_3,\ x_8 R x_4,\ x_3 R x_8,\ x_4 R x_7,\ x_{11} R x_{12},\ x_{12} R x_1\ x_{12} R x_2,\ x_9 R x_5,\ \text{ and } x_{10} R x_6,$$

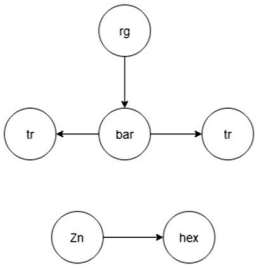
which generates the graph $G_{NS-III}(X, L, R)$ shown in Fig. 9b. In graph $G_{NS-III}(X, L, R)$, as in $G_{NS-III}(X, L, R)$, no two nodes that have the same label can be distinguished. Thus, the partition of the set of nodes has the form (2, 2, 2, 2, 2, 1, 1) and

$$H^{nodes}(\text{ NS-III}) = 33.007.$$

In $G_{NS-III}(X, L, R)$ there are four pairs of indistinguishable edges: (arm , Cu), (Cu , arm), (bar , tr) and (Zn , hex). Furthermore, there is a single distinguishable edge—(rg , bar). This creates in the set of edges the partition (2, 2, 2, 2, 1)(132), so







(a) the created structures **(b)** relations

Fig. 9 NetState III (NS-III)

$$H^{edges}(\text{ NS-III}) = -9 \cdot \left[4 \cdot \left(\frac{2}{132} \cdot \log \frac{2}{132}\right) + \left(\frac{1}{132} \cdot \log \frac{1}{132}\right)\right] = 3.737$$

and

$$H(\text{ NS-III}) = 36.744.$$

Let us discuss the obtained results. Intuitively, according to the complexity of the structures (see Figs. 5, 8, 9), we expect that the amount of information should be significantly less in NS-I than in both NS-II and NS-III. Furthermore, the amount of information in NS-II and NS-III should be approximately equal. The last condition is satisfied. The first condition, however, is satisfied only partially. It is true that the amount of information in NS-II and NS-III is greater than in NS-I, but not much. Contrary to appearances, however, in the case under consideration, this result is correct. A large amount of information H^L generated by labeling the elements of the set X, that is common to all three states, is the reason for the small difference in the amount of information between NS-I and NS-II as well as between NS-I and NS-III. Labeling generates partition (2, 2, 2, 2, 2, 1, 1) of X which corresponds to generating large amount of information. This causes the additional node information generated by the structure to be zero (NS-I and NS-III) or small (NS-II). On the other hand, the amount of edge information H^{edges} is small in comparison with node information H^{node} because in all three states the number of existing edges of the graph is small in comparison with the number of edges in the fully connected graph.

In support of the above claims, let us perform calculations for the above structures assuming no labeling. In NS-I, see Fig. 5b, there are four indistinguishable nodes cyclically connected by edges and eight isolated nodes, which generates the partition of the nodes (4, 8). The partition of the edges is (4)(132). Thus

$$H^{nodes}(\text{ NS-I}) = -12 \cdot \left[\left(\frac{4}{12} \cdot \log \frac{4}{12}\right) + \left(\frac{8}{12} \cdot \log \frac{8}{12}\right)\right] = 11.016.$$

and

$$H^{edges}(\text{ NS-I}) = -4 \cdot \left[\left(\frac{4}{132} \cdot \log \frac{4}{132}\right)\right] = 0.611.$$

As a result

$$H(\text{ NS-I}) = 11.627.$$

In NS-II, see Fig. 8b, there are three isolated nodes, four indistinguishable nodes in the cyclic component of the graph, and five nodes in the acyclic component of the graph. In the last subset of the nodes, three nodes are fully distinguishable—one has only outcoming edges, one has two outcoming edges and the third one has one outcoming edge and one incoming edge. The other two vertices only have one outcoming edge, so they are distinguishable from the other three, but indistinguishable from each other. To sum up, unlabeled graph in Fig. 8b generates the partition (4, 3, 2, 1, 1, 1) in the set of the nodes. Similarly, in the set of the graph edges, the partition (4, 2, 1, 1)(132) is generated. Thus

$$H^{nodes}(\text{ NS-II}) = -12 \cdot \left[\left(\frac{4}{12} \cdot \log \frac{4}{12}\right) + \left(\frac{3}{12} \cdot \log \frac{3}{12}\right) + \left(\frac{2}{12} \cdot \log \frac{2}{12}\right) + 3 \cdot \left(\frac{1}{12} \cdot \log \frac{1}{12}\right)\right] = 28.260.$$

and

$$H^{edges}(\text{ NS-II}) = -8 \cdot \left[\left(\frac{4}{132} \cdot \log \frac{4}{132}\right) + \left(\frac{2}{132} \cdot \log \frac{2}{132}\right) + 2 \cdot \left(\frac{1}{132} \cdot \log \frac{1}{132}\right)\right] = 2.810.$$

As a result

$$H(\text{ NS-II}) = 31.070.$$

In NS-III, see Fig. 9b, there are four connected components of the graph. Two of them consist of two nodes connected by one edge. Two nodes of these components with outcoming edges are indistinguishable, as well as two nodes with incoming edges. Two additional indistinguishable nodes, each with one incoming edge, exist in the non-cyclic component that consists of four edges. Two other nodes in this component are fully distinguishable. Thus, we have three pairs of indistinguishable nodes, two fully distinguishable nodes, and four indistinguishable nodes in the cyclic component which gives partition (4, 2, 2, 2, 1, 1) and the following amount of node information is generated

$$H^{nodes}(\text{ NS-III}) = -12 \cdot \left[\left(\frac{4}{12} \cdot \log \frac{4}{12}\right) + 3 \cdot \left(\frac{2}{12} \cdot \log \frac{2}{12}\right) + 2 \cdot \left(\frac{1}{12} \cdot \log \frac{1}{12}\right)\right] = 29.014.$$

In the set of edges, there are four indistinguishable edges in the cyclic component and one fully distinguishable edge that outcomes from the node that has not any incoming edge in the acyclic component that consists of four nodes. The other two vertices in the acyclic four-node component are indistinguishable from each other as well as the two in the components that consist of two nodes. This generates partition (4, 2, 2, 1) and the following amount of information,

$$H^{edges}(\text{ NS-II}) = -9 \cdot \left[\left(\frac{4}{132} \cdot \log \frac{4}{132}\right) + 2 \cdot \left(\frac{2}{132} \cdot \log \frac{2}{132}\right) + \cdot \left(\frac{1}{132} \cdot \log \frac{1}{132}\right)\right] = 3.503.$$

As a result

$$H(\text{ NS-III}) = 32.517.$$

The calculations done for the unlabeled version confirm the hypothesis that in the considered case the labeling provides a large amount of information that for NS-I and NS-III fully and for NS-II largely overlaps the structural node information. This causes the mentioned small difference between the amount of information between NS-I and both NS-II and NS-III for the labeled version.

5 Discussion

A question naturally arises as to the usefulness of the proposed theoretical concept. It should be stressed strongly, that in this section we will think beyond static cases.

Let us start with the idea of molecular robots in which all modules, i.e. not only sensors and effectors but also the controlling computers are molecular devices, as outlined by Hagiya et al. (2016). As in these molecular computers the authors only considered DNA computing. They pointed out that this type of molecular computing does not satisfy reusability which means that the computational device can be used only once. Such an approach implies, in turn, the necessity of a continuous supply of new devices—(Hagiya et al., 2016), Section 2. In the light of the concept presented in our paper, however, computations

do not need to be performed digitally, i.e. on the basis of processing digital code such as DNA. Computations may be performed based on creating new structures and modifying existing ones as well as modifying relations between structures. This opens up a new perspective on the problem of information processing in general and the problem of computing in particular. The problem is discussed in detail in Bielecki (2015), section 5.1, where prokaryotic transcription is studied in this context. It should be also mentioned that the problem of biological and molecular computations and the possibility of their applications is a topic of some studies—see, for instance, (Burgin & Adamatzky, 2017) and references given there.

The theory proposed in our paper enables us to quantify the information encoded in the structures and processed by altering the structures. The change in the value of H informs about the changing amount of information encoded in a structural transformation, a protocol that can be readily applied to characterized molecular computers. In this type of computer, computations will be performed on the basis of changes in molecular structures. As a consequence, data will have to be encoded in molecular structures, as well.

The problem of autonomous systems strongly refers to robotics and, consequently, to molecular cybernetics. In the years to come, autonomous systems will face numerous new types of challenges which, in turn, will enforce a new approach to the design, analysis, and construction of autonomous robots (Harel et al., 2020). In this type of robot, the issue of their decomposability is crucial. It enables to understand and anticipate their behavior by decomposition of the whole autonomous system into well-understood functional and structural subsystems. The presented theory provides additional tools for the required analysis, making it possible to study the degree of complexity of the robot which, in turn, determines its computational capabilities. This information is crucial especially when constructing the above-discussed molecular robots, the complexity of which is related to its computational capabilities in a very direct way.

In biology, our approach will provide a tool for analyzing information processing on various levels of complexity. First of all, it was postulated that three-componential processing, in which information, matter and energy are closely intertwined, is the basis of metabolism in living individuals—see Bielecki (2015) and references given there. At the subcellular level, for instance, the interaction between cell receptor and ligand—that trigger a signal pathway—is an example of information processing realized by modifying structures with simultaneous processing of energy. The aforementioned symbiosis and endosymbiosis consist in connecting two structures. The natural question emerges whether the amount of information encoded in the created structure is greater, equal or less than the sum of information contained in the components. The proposed approach can be also used for analyzing the hypothetical creation of information during transitions in evolution (Rosslenbroich, 2014).

To sum up, the proposed formal approach can be applied in molecular robotics to assess the computational power of the robot modules, for instance, receptors but, first of all, molecular computers. Furthermore, based on the introduced theory, molecular devices may be designed precisely according to their informational capacity. Various models of cooperation between molecular robots, including their structural and functional combination (symbiotic computing—(Harel et al., 2020)), which corresponds to biological symbiosis and endosymbiosis, can be analyzed with regard to the information processing ability of the robot. Additionally, the proposed formalism will allow the evaluation of informational aspects of life processes, which, on the one hand, have so far been less studied than structural and energetic aspects and, on the other hand, belong to the fundamental manifestations of life.

6 Concluding Remarks

The formalization introduced in Sect. 2 made it possible both to define various aspects of information related to the structure and to measure its quantity. The formalization is based on the graph generated by the elements and relations between them. The proposed approach has been applied in molecular cybernetics to measure the amount of information in the structure of molecular machines. This is the first step towards applying it to biological structures.

As it has been already mentioned, the formalization presented in this article refers only to the mathematical aspects of the theory of structural information adequate in biological applications and covers only the case of static, non-hierarchical structures. The planned subsequent steps will include:

- Information in hierarchical structures,
- Information in dynamical structures, including self-replicate systems
- Extending the approach to the physical aspect of structured information, including energetic aspects.

Let us explicitly emphasize that the aim of the undertaken studies is to enable the investigation of biological processes in their informative aspect. The biological entities are structures that, in a way, compute themselves by, among others, modifying their structures. Therefore, working out the approach enabling analysis of the informative aspect of structures as such is a necessary step to attack the problem effectively. It should be also ephasized that the presented version of information theory can be also used in network structures, including artificial ones, for which adequate information theory is also look for—see, for instance, (Tao et al., 2021).

Funding The paper was not support by any funding.

Declarations

Conflict of interest The author declare no conflict of interests.

References

Baroncini, M., Silvi, S., & Credi, A. (2020). Photo- and redox-driven artificial molecular motors. *Chemical Reviews, 120,* 200–268.

Barreiro, C., Barreiro, J. M., Lara, J. A., Lizcano, D., Martínez, M. A., & Pazos, J. (2020). The third construct of the Universe: Information. *Foundations of Science, 25,* 425–440.

Bielecki, A. (2015). A general entity of life: A cybernetic approach. *Biological Cybernetics, 109,* 401–419.
Bielecki A. (2019). Models of Neurons and Perceptrons: Selected Problems and Challenges, Springer, series: Studies in Computational Intelligence.
Biswas, P. K., Saha, S., Gaikwad, S., & Schmittel, M. (2020). Reversible multicomponent AND gate triggered by stoichiometric chemical pulses commands the self-assembly and actuation of catalytic machinery. *Journal of the American Chemical Society, 142,* 7889–7897.
Boyer, P. D. (1998). Energy, Life, and ATP (Nobel Lecture). *Angewandte Chemie Int. Ed., 37,* 2296–2307.
Burgin, M., & Adamatzky, A. (2017). Structural machines as a mathematical model of biological and chemical computations. *Theory and Applications of Mathematics and Computer Science, 7*(2), 1–30.
Carnap, R. (1928). *Der logische Aufbau der Welt*. Meiner Verlag.
Goldfeld, Z., Bresler, G., & Polyanskiy, Y. (2021). Information storage in the stochastic Ising model. *IEEE Transactions on Information Theory, 67,* 1373–1399.
Goswami, A., Paululat, T., & Schmittel, M. (2019). Switching dual catalysis without molecular switch: Using a multicomponent information system for reversible reconfiguration of catalytic machinery. *Journal of the American Chemical Society, 141,* 15656–15663.
Hagiya, M., Aubert-Kato, N., Wang, S., & Kobayashi, S. (2016). Molecular computers for molecular robots as hybrid systems. *Theoretical Computer Science, 632,* 4–20.
Harel, D., Marron, A., & Sifakis, J. (2020). Autonomics: In search of a foundation for next generation autonomous systems. *Proceedings of the National Academy of Sciences, 117,* 17491–17498.
Hellerman, L. (2006). Representation of living forms. *Biology and Philosophy, 21,* 537–552.
Hellerman, L. (2016). The animate: Inanimate relationship. *International Journal of General Systems, 45,* 734–746.
Jablonka, E., & Lamb, M. J. (2006). The evolution of information in the major transitions. *Journal of Theoretical Biology, 239,* 236–246.
Krzanowski, R. (2020). What is physical information? *Philosophies, 5,* 1–19.
Krzanowski, R. (2020). Why can information not be defined as being purely epistemic? *Zagadnienia Filozoficzne w Nauce (Philosophical Problems in Science), 68,* 37–62.
Nurse, P. (2008). Life, logic and informaton. *Nature, 454,* 424–426.
Paul, I., Ghosh, A., Bolte, M., & Schmittel, M. (2019). Remote control of the synthesis of a [2]rotaxane and its shuttling via metal-ion translocation. *ChemistryOpen, 8,* 1355–1360.
Perez Velazquez, J. L. (2005). Brain, behaviour and mathematics: Are we using the right approaches? *Physica D, 212,* 161–182.
Perez Velazquez, J. L. (2009). Finding simplicity in complexity: General principles of biological and non-biological organization. *Journal of Biological Physics, 35,* 209–221.
Qiu, Y., Feng, Y., Guo, Q. H., Astumian, R. D., & Stoddart, J. F. (2020). Pumps through the ages. *Chem, 6,* 1952–1977.
Rodnina, M. V., & Wintenmeyer, W. (2011). The ribosome as a molecular machine: The mechanism of tRNA-mRNA movement in translocation. *Biochemical Society Transactions, 39,* 658–662.
Rohwer, F., & Barott, K. (2013). Viral information. *Biology and Philosophy, 28,* 283–197.
Rosslenbroich B. (2014). On the Origin of Autonomy. *A New Look at the Major Transitions in Evolution,* Springer, series: History, Philosophy and Theory of the Life Sciences (vol. 5).
Sadeh, S., & Clopath, C. (2020). Theory of neuronal perturbome in cortical networks. *Proceedings of the National Academy of Sciences, 117*(43), 26966–26976.
Schmittel, M. (2019). Dynamic functional molecular systems: From supramolecular structures to multicomponent machinery and to molecular cybernetics. *Israel Journal of Chemistry, 59,* 197–208.
Schmittel, M., & Howlader, P. (2020). Toward molecular cybernetics: The art of communicating chemical systems. *Chemical Record, 20,* 1–22.
Schepherd, G. M. (1994). *Neurobiology*. Oxford University Press.
Schneider, T. D. (2000). Evolution of biological information. *Nucleid Acids Research, 28,* 2794–2799.
Smith, J. M. (2000). The concept of information in biology. *Philosophy of Science, 67,* 177–194.
Spirin, A. S. (2002). Ribosome as a molecular machine. *FEBS Letters, 514,* 2–10.
Spirin, A. S. (2009). The ribosome as a conveying thermal ratchet machine. *Journal of Biological Chemistry, 284,* 21103–21119.
Tadeusiewicz, R. (2010). New trends in neurocybernetics. *Computer Methods in Material Science, 10*(1), 1–7.
Tao, M., Wang, S., Chen, H., & Wang, X. (2021). Information space of multi-sensor networks. *Information Sciences, 565,* 128–145.
van Leeuwen, T., Lubbe, A. S., Stacko, P., Wezenberg, S. J., & Feringa, B. L. (2017). Dynamic control of function by light-driven molecular motors. *Nature Reviews Chemistry, 1,* 96.

Walker, S. I., Kim, H., & Davies, P. C. W. (2016). The information architecture of the cell. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374,* 1–20.

Andrzej Bielecki received the M.Sc. degree in Physics and Mathematics from the Jagiellonian University in 1985 and 1992 respectively, Ph.D. in Mathematics in 1999, D.Sc. (habilitation) in Mathematics in 2009 and professorship in 2020 in Computer Science. Dynamical systems theory, artificial intelligence, cybernetics and philosophy of science are the topics of his scientific interest. He is an author of over 120 scientific papers and one textbook.

Michael Schmittel studied chemistry and French in Freiburg and Paris. After habilitation at the University of Freiburg in 1992, he joined the University of Würzburg as an associate professor in 1993 and since 1999 holds a Chair in Organic Chemistry at the University of Siegen. His broad interests, published in more than 320 peer-reviewed papers and one textbook, cover research areas as diverse as radical cation catalysis, nonstatistical dynamics, thermal diradical cyclizations, metallosupramolecular coordination chemistry, multichannel chemosensors and, more recently, (supra)molecular machinery and nanoswitches for catalysis.