A formal description of zz-structures

Antonina Dattolo
Dipartimento di Matematica e Informatica
Università degli Studi di Udine, Italy
antonina.dattolo@uniud.it

Flaminia L. Luccio
Dipartimento di Informatica
Università Ca’ Foscari, Venezia, Italy
luccio@unive.it

ABSTRACT
The focus of this paper is on particular and innovative structures for storing, linking and manipulating information: the zz-structures.

In the last years, we worked at the formalization of these structures, retaining that the description of the formal aspects can provide a better understanding of them, and can also stimulate new ideas, projects and research.

This work presents our contribution for a deeper discussion on zz-structures.

1. INTRODUCTION
At the first Wearable Computer Conference [18], Ted Nelson proposed “a prototype that implements some interesting ideas, intended to lead to such a new kind of simple and unified world, possibly to permit the unification of everything that non-computer people want to do with computers”. The software was ZigZag™, a “principled system of interconnections” [19] and a new, graph-centric system of conventions for data and computing, based on the so-called zz-structures.

Nelson writes [19]: “The ZigZag system is very hard to explain, especially since it resembles nothing else in the computer field that we know of, except perhaps a spreadsheet cut into strips and glued into loops”. Zz-structures are hyper-orthogonal, non-hierarchical structures for storing, linking and manipulating data. Intuitively, data are contained in cells that are connected by crossing dimensions forming a structure that resembles a spreadsheet but contains intricate connections [18, 19, 20, 21, 22, 23].

Zz-structures have been seen from many different perspectives: merely a large variety of implementations and applications, and some formal models have been proposed. We specify that we use the term “applitude” instead of “application”. Nelson states [23]: “Instead of “applications”, separated zones of function and usage connected by the narrow channels of clipboard and file export/import, we have “applitudes” which are deeply interconnected to the whole, amongst themselves, and amongst their parts”. The main difference is that, unlike applications, applitudes exploit the dense and intricate connections among the information contained in a zz-structure. Applitudes can also be combined with each other and are not “walled off” from the rest of the system. Table 1 collects a list of research contributions (applitudes and implementations), developed using zz-structures.

Zz-structures are “a generalized representation for all data and a new set of mechanisms for all computing” [23]: innovative structures for storing, linking and manipulating information.

The intention of this work is to present our contribution to the formalization of zz-structures and to encourage further discussion. In our opinion, the description of a formal model can provide a deeper understanding of the model and can stimulate new ideas, projects and research.

The paper is organized as follows: in Section 2, we give an informal description of zz-structures, that prepares the reader to the formal model presented in Section 3. We conclude in Section 4 with a brief discussion and future look.

2. THE ZZ-STRUCTURES
Zz-structures introduce a new, graph-centric system of conventions for data and computing [19]. A zz-structure can be thought of as a space filled with cells. Each cell may have a content (such as integers, text, images, audio, etc.), and it is called atomic if it contains only one unit of data of one type, or it is called referential if it represents a package of different cells [23].

Cells are connected together with links of the same color into linear sequences called dimensions. A single series of cells connected in the same dimension is called rank, i.e., a rank is in a particular dimension and a dimension may contain many different ranks. The starting and the ending

Table 1: Models and applitudes

<table>
<thead>
<tr>
<th>Year</th>
<th>Domain</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984-2004</td>
<td>zz-vims (Azz, Ezz, Gzz and Zzz, and Lzz)</td>
<td>[23]</td>
</tr>
<tr>
<td>2001</td>
<td>Associative writing</td>
<td>[24]</td>
</tr>
<tr>
<td>2001</td>
<td>Various Demos</td>
<td>[4]</td>
</tr>
<tr>
<td>2002-2006</td>
<td>Bioinformatics</td>
<td>[16, 17]</td>
</tr>
<tr>
<td>2004</td>
<td>Cellular phones</td>
<td>[17]</td>
</tr>
<tr>
<td>2006-2007</td>
<td>Archival finding aids</td>
<td>[1]</td>
</tr>
<tr>
<td>2007</td>
<td>Grid Systems</td>
<td>[9]</td>
</tr>
<tr>
<td>2007-2009</td>
<td>Web-based Education</td>
<td>[2, 10, 12]</td>
</tr>
<tr>
<td>2007-2009</td>
<td>Audio Archives</td>
<td>[5, 7]</td>
</tr>
<tr>
<td>2008</td>
<td>Personal Information Space</td>
<td>[5]</td>
</tr>
<tr>
<td>2008</td>
<td>Sentiment Classification</td>
<td>[6]</td>
</tr>
<tr>
<td>2009</td>
<td>Publication Sharing Systems</td>
<td>[8]</td>
</tr>
</tbody>
</table>
cells of a rank are called, headcell and tailcell, respectively, and the direction from the starting (ending) to the ending (starting) cell is called posward (respectively, negward). For any dimension, a cell can only have one connection in the posward direction, and one in the negward direction. This ensures that all paths are non-branching, and thus embodies the simplest possible mechanism for traversing links. Dimensions are used to project different structures: Ordinary lists are viewed in one dimension; spreadsheets and hierarchical directories in many dimensions.

There are many different ways to view these structures: A raster is a way of selecting the cells from a structure, while a view is a way of placing the cells on a screen. Generic views are designed to be used in a big variety of cases and usually show only few dimensions or few steps in each dimension. Among them the most common views are the two-dimensional rectangular views: The cells are placed, using different rasters, on a Cartesian plane where the dimensions increase going down and to the right. The simplest raster is the row and column raster, i.e., two rasters which are the same but rotated of 90 degrees from each other. A cell is chosen and placed at the center of the plane (cursor centric view). The chosen cell, called focus, may be changed by moving the cursor horizontally and vertically. In a row view \( i \), a rank is chosen and placed vertically. Then the ranks related to the cells in the vertical rank are placed horizontally. Vice versa, in the column view \( j \), a rank is chosen and placed horizontally and the related ranks are placed vertically. All the cells are denoted by different numbers. Note that in a view the same cell may appear in different positions as it may represent the intersection of different dimensions.

3. **FORMALIZING ZZ-STRUCTURES**

As we have mentioned in the introduction, zz-structures have been studied from many perspectives, and many implementations and applications have been proposed in different fields. Some research has also been provided towards a formal definition of these structures. In the seminal work of [14], we find the first proposal of formalizing these structures in terms of graphs; this work has been extended and motivated in [15], where the authors use this formalization in order to compare these structures with mSpaces and Polymarchies. This comparison is done building a taxonomy as a subsumption diagram, a subsumption being a generalization of something. The general result is that zz-structures subsume lists, 2D arrays, trees and also polyarchies; polymarchies subsume mSpace polymarchies. Finally zz-structures and edge-colored multigraph subsume each other.

Later, in [9, 10, 11, 12] we have revisited and redefined into more precise mathematical terms, the definitions of zz-structures provided in [14, 15], and have also introduced new notions such as the one of local orientation, borrowed from the field of distributed computing, and required to provide a formal definition of local posward and negward directions. The formal model is a requirement in [9] for the construction of an actor-based model where actors can add new connections between cells of the zz-structure, i.e., dynamically modifying its structure. Moreover, this formalization has provided interesting tools for the introduction of new formal concepts, such as the extension of the standard notion of view to higher dimensional views (e.g., \( n \)-dimensions H and I views, 3-dimensions extended H and I views, etc.) [10, 12], and for the definition of techniques that allow users to display neighbouring views (i.e., views centered in a cell at distance one from the previous one) [11].

The aim of this section is to summarize, analyze, and illustrate with new examples, discuss and relate all the above proposals in order to provide a general overview of the formal model, at least for the original concepts defined by Nelson in [18].

**Zz-structure.**

In [14, 15] the authors define a zz-structure as a directed multigraph with colored edges (i.e., a graph where pair of nodes may have multiple colored edges connecting them), where each node has at most one outgoing edge and one outgoing edge for each color. In [9] the authors further formalize these concepts (choosing however bidirectional links) as follows.

Consider an edge-colored multigraph \( \text{ECMG} = (MG, C, c) \) where: \( MG = (V, E, f) \) is a multigraph composed of a set of vertices \( V \), a set of edges \( E \) and a surjective function \( f : E \rightarrow \{u, v\} \mid u, v \in V, u \neq v \). \( C \) is a set of colors, and \( c : E \rightarrow C \) is an assignment of colors to edges of the multigraph. Finally, \( \text{deg}(x) \) (respectively, \( \text{deg}_c(x) \)) denotes the number of edges incident to \( x \), (respectively, of color \( c_0 \)).

*Proposition 1.* Consider a set of colors \( C = \{c_1, c_2, \ldots, c_{|C|}\} \) and a family of indirect edge-colored graphs \( \{D^1, D^2, \ldots, D^{|C|}\} \), where \( D^k = (V, E^k, f, \{c_k\}, c) \), with \( k = 1, \ldots, |C| \), is a graph such that: 1) \( E^k \neq \emptyset \); 2) \( \forall x \in V, \text{deg}_k(x) = 0, 1, 2 \). Each vertex of a zz-structure is called zz-cell and each edge a zz-link. The set of isolated vertices is \( V_0 = \{x \in V : \text{deg}_k(x) = 0\} \).

An example of a zz-structure is shown in Figure 1. Normal, dotted and thick lines represent different colors.

**Dimension.**

In [15] the authors state that “each of the edge colors correspond to a different spatial dimension”. This concept, together with all the following definitions, is further formalized in [9] where the authors state that an alternative way of viewing a zz-structure is a union of subgraphs, each of them containing edges of a unique color.

*Proposition 1.* Consider a set of colors \( C = \{c_1, c_2, \ldots, c_{|C|}\} \) and a family of indirect edge-colored graphs \( \{D^1, D^2, \ldots, D^{|C|}\} \), where \( D^k = (V, E^k, f, \{c_k\}, c) \), with \( k = 1, \ldots, |C| \), is a graph such that: 1) \( E^k \neq \emptyset \); 2) \( \forall x \in V, \text{deg}_k(x) = 0, 1, 2 \). Then, \( S = \bigcup_{k=1}^{|C|} D^k \) is a zz-structure.

*Definition 2.* Dimension - Given a zz-structure \( S = \bigcup_{k=1}^{|C|} D^k \), then each graph \( D^k, k = 1, \ldots, |C| \), is a distinct dimension of \( S \).
The zz-structure of Figure 1 contains three dimensions $D^{\text{thick}}$, $D^{\text{normal}}$ and $D^{\text{dotted}}$, respectively represented by thick, normal, and dotted lines and shown in Figure 2. In turn, each dimension is composed by a set of connected components and a set (eventually empty) of isolated vertices. As an example, $D^{\text{normal}}$ is composed of a cycle $\{v_1, v_2, v_3, v_4\}$, a path $\{v_3, v_5\}$, and one isolated vertex $v_6$, while $D^{\text{thick}}$ is composed of two distinct paths $\{v_2, v_3, v_4, v_5, v_6\}$, and no isolated vertex.

**Rank.**

Each “series of cells connected sequentially in any dimension” identifies a rank [23].

**Definition 3.** Rank - Consider a dimension $D^k = (V, E^k, f, \{c_k\}, c)$, $k = 1, \ldots, |C|$ of a zz-structure $S = \bigcup_{k=1}^{|C|} D^k$. Then, each of the $l_k$ ($l_k \geq 1$) connected components of $D^k$ is called a rank.

Thus a rank is an indirect graph $R^k = (V^k, E^k, f, \{c_k\}, c)$ ($i = 1, 2, \ldots, l_k$) such that 1) $E^k \subseteq E^k$ and $E^k \neq \emptyset$; 2) $\forall x \in V^k, V^k \subseteq V \Leftrightarrow \deg_k(\{x\}) = 1, 2$.

**Definition 4.** Ringrank - A ringrank is a rank $R^k$, where $\forall x \in V^k, \deg_k(\{x\}) = 2$.

In Figure 2, the dimension $D^{\text{thick}}$ has two ranks: $\{v_3, v_1\}$ and $\{v_2, v_3, v_4, v_5\}$; the dimension $D^{\text{normal}}$ has one rank $\{v_3, v_5\}$, and one ringrank $\{v_1, v_2, v_4, v_1\}$.

**Cells and their orientation.**

A vertex [9] has local orientation on a rank if each of its (1 or 2) incident edges has assigned a distinct label (1 or -1). More formally (see also [13]):

**Definition 5.** Local orientation - Consider a rank $R^k = (V^k, E^k, f, \{c_k\}, c)$ of a zz-structure $S = \bigcup_{k=1}^{|C|} D^k$. Then, $\exists g^k: E^k \to \{-1, 1\}$ such that, $\forall x \in V^k, g^k(x, y) \in E^k$, if $\exists y, z \in V^k : \{x, y\}, \{x, z\} \in E^k$, then $g^k(\{x, y\}) \neq g^k(\{x, z\})$. Thus, we say that each vertex $x \in V^k$ has a local orientation in $R^k$.

**Definition 6.** Posward and negward directions - Given an edge $[a, b] \in E^k$, we say that $[a, b]$ is in a posward direction from $a$ in $R^k$, and that $b$ is its posward cell if $g^k([a, b]) = 1$, else $[a, b]$ is in a negward direction and $a$ is its negward cell. Moreover, a path in rank $R^k$ follows a posward (negward) direction if it is composed of a sequence of edges of value 1 (respectively, -1).

**Head and tail cells.**

If we focus on a vertex $x$, $R^k = \ldots x^{-2}x^{-1}x^{1}x^{2} \ldots$ is expressed in terms of negward and posward cells of $x$: $x^{-1}$ is the negward cell of $x$ and $x^{1}$ the posward cell. We also assume $x^0 = x$. In general $x^{-i}(x^{i+1})$ is a cell at distance $i$ in the negward (posward) direction.

**Definition 7.** Headcell and tailcell - Given a rank $R^k = (V^k, E^k, f, \{c_k\}, c)$, a cell $x$ is the headcell of $R^k$ iff $\exists x \subseteq x^{1}$ and $x^{2}$ its negward cell $x^{-1}$. Analogously, a cell $x$ is the tailcell of $R^k$ iff $\exists x \subseteq x^{-1}$ and $x^{1}$ its posward cell $x^{1}$.

**Views.**

In the following, we denote with $x \in R^k(x)$ the rank $R^k(x)$ related to vertex $x$, of color $c_o$.

**Definition 8.** H-view - Given a zz-structure $S = \bigcup_{k=1}^{|C|} D^k$, where $D^k = \bigcup_{l=1}^{|l_{k}|} D^l$, and where $R^k = (V^k, E^k, f, \{c_k\}, c)$, the H-view of size $l = 2m + 1$ and of focus $x \in V = \bigcup_{l=1}^{|l_{k}|} V^l$, on main vertical dimension $D^y$ and secondary horizontal dimension $D^x$ is:

- the central node, in position $((m + 1), (m + 1))$, is the focus $x$;
- the horizontal central path (the $m$-th row) from left to right, focused in vertex $x \in R^k(x)$ is: $x^{-g} \ldots x^{-1}x^{1}x^{2} \ldots$; where $x^s \in R^k(x)$ for $s = -g, \ldots, +p$ ($p \leq m$).
- for each cell $x^{s}$, $s = -g, \ldots, +p$, the related vertical path, from top to bottom, is: $(x^{s})^{-g} \ldots (x^{s})^{-1}x^{m} \ldots (x^{s})^{p}$, where $(x^{s})^{m} \in R^k(x^{m}), g \leq m$.

Intuitively, the H-view extracts ranks along the two chosen dimensions. Note that, the name H-view comes from the fact that the columns remind the vertical bars in a capital letter H. Observe also that the cell $x^{-g}$ (in the $m$-th horizontal row) is the headcell of $R^k(x)$, if $g < m$ and the cell $x^{m}$ (in the same row) is the tailcell of $R^k(x)$, if $p < m$. Analogously, the cell $x^{-p}$ is the headcell of $R^k(x)$, if $g > m$ and the cell $x^{m}$ is the tailcell of $R^k(x)$, if $p > m$. Intuitively, the view is composed of $l \times l$ cells unless some of the displayed ranks have their headcell or tailcell very close (less than $m$ steps) to the chosen focus.

As an example consider Figure 3 left that refers to the zz-structure of Figure 1. The main vertical dimension is $D^{\text{dotted}}$ and the secondary horizontal dimension is $D^{\text{thick}}$. The view has size $l = 2m + 1 = 5$, the focus is the node $v_5$, the horizontal central path is $\{v_2, v_3, v_5, v_6\}$, the vertical path related to $v_4$ is $\{v_3, v_4, v_5, v_6\}$, that is $v_5$ is the tailcell of the rank as $p_5 = 1 < m = 2$.

The I-view can be defined analogously to the H-view [9]. An example of I-view with main horizontal dimension $D^{\text{dotted}}$, secondary vertical dimension $D^{\text{thick}}$, size $l = 5$ and focus $v_3$ is shown in Figure 3 right.

We can now extend the known definition of H and I views to a number $n > 2$ of dimensions [10]. Intuitively, we will
build $n-1$ different $H$-views (respectively, $I$-views), centered in the same focus, with a fixed main dimension and a secondary dimension chosen among the other $n-1$ dimensions. Formally:

**Definition 9.** **n-dimensions H-view** - Given a $zz$-structure $S = \bigcup_{i=1}^{[n]} D^{k}_i$, where $D^{k}_i = \bigcup_{x=1}^{[m]} (R^{c}_i \cup V^{k}_i)$, and where $R^{c}_i = (V^{c}_i, E^{c}_i, f, \{c_k\}, c)$, the $n$-dimensions $H$-view of size $l = 2m + 1$ and of focus $x$, $x \in V = \bigcup_{k=0}^{[n]} V^{k}$, on dimensions $D^{1}, D^{2}, \ldots, D^{n}$ is composed of $n-1$ rectangular $H$-views, of main dimension $D^{i}$ and secondary dimensions $D^{j}$, $i = 2, \ldots, n$, all centered in the same focus $x$.

Analogously, we can define an $n$-dimensions $I$-view. An example of a 3-dimensions $H$-view is provided in Figure 4.

![Figure 3: An example of $H$-view and $I$-view.](image)

**Figure 4: An example of a 3-dimensions $H$-view.**

This view has focus on $v_3$, size $l = 5$, main dimension $D_{\text{dotted}}$, and secondary dimensions $D_{\text{dotted}}$, and $D_{\text{dotted}}$.

A star view [12] visualizes information related to a focus vertex and a set of $n$ chosen dimensions. There are two typologies of star views: the star view and the $m$-extended star view.

**Definition 10. Star view** - Given a $zz$-structure $S = \bigcup_{i=1}^{[c]} D^{k}_i$, where $D^{k}_i = \bigcup_{x=1}^{[m]} (R^{c}_i \cup V^{k}_i)$, and where $R^{c}_i = (V^{c}_i, E^{c}_i, f, \{c_k\}, c)$, the star view of focus $x \in V = \bigcup_{k=0}^{[n]} V^{k}$ and dimensions $D^{1}, D^{2}, \ldots, D^{n}$ is a star graph $n_{+1}$-star on central vertex $x$ and neighborhood $N(x) = \{y \in V : y = x^{+1}, x^{+1} \in R^{c}_i, i \in \{1, \ldots, n\}\}$.

The $m$-extended star view extends the number of cells directly accessible from a view; it is based on a $star$ view, but, for each vertex $y$ in the neighborhood $N(x)$, adds the set of the $p (p \leq m)$ postward cells related to the given dimensions.

**Definition 11. $m$-extended star view** - Given a $zz$-structure $S = \bigcup_{i=1}^{[c]} D^{k}_i$, where $D^{k}_i = \bigcup_{x=1}^{[m]} (R^{c}_i \cup V^{k}_i)$, and where $R^{c}_i = (V^{c}_i, E^{c}_i, f, \{c_k\}, c)$, the $m$-extended star view is a star view of focus $x \in V = \bigcup_{k=0}^{[n]} V^{k}$, dimensions $D^{1}, D^{2}, \ldots, D^{n}$, and each extension constituted, $vy \in N(x)$ and $\forall i \in \{1, \ldots, n\}$, by the paths $(y^{+1}, \ldots, y^{+p}) \subseteq R^{c}_i (p \leq m)$.

Fig. 5 shows an schematic example of 5-extended star view.

![Figure 5: A 5-extended star view.](image)

The view has focus $v_3$ and shows the connections along six dimensions. Note that cells $v_1$ and $v_7$ have, e.g., extension $p = 0$, while the maximum extension $(p = m = 5)$ is reached only by the connection along the dashed and double lines.

For lack of space, we cannot provide all the formal extensions to multidimensional views, the algorithms for the additions of new connections and the displaying of neighboring views and we refer interested readers to [9], [10], [11], [12]. Just note that the above formal models have found interesting real world applications. Very briefly, in [9] the authors present an actor-based model, capable of representing both hypermedia distribution and collaborative schemes among different and heterogeneous entities which are part of a particular grid infrastructure and cooperate in order to achieve common goals and solve problems. In [10, 12], the authors propose the use $zz$-structures in order to help an author of an e-learning environment, to organize documents on a given topic in a concept space, and to create semantic interconnections and personalized maps. Finally, in [11] the authors propose a multi-agent adaptive system to support tours of virtual museums. In particular, the agents collaborate in order to help users visualizing their personalized views and choosing their navigational path inside the virtual museum.

4. DISCUSSION AND CONCLUSIONS

In this paper we have concentrated our attention on the formal models for representing $zz$-structures. Besides the motivations, above provided, in our opinion a formal model can help the navigation of a user by providing extra information, such as, e.g., the distances between the cell where (s)he
is located and one where (s)he wants to move. Defining zz-structures as graphs allows, e.g., the application of known algorithms for the (dynamical) computation of shortest paths, or of paths with small stretch factor (i.e., ratio between the best path connecting two nodes and the shortest path). A user may thus compute a general shortest path, or, e.g., a shortest path (given that the two nodes are connected) in the subgraph induced by a particular color, meaning that (s)he wants to move following a unique dimension, i.e., concept.

Currently we are extending the formal model and preparing a survey of current literature on zz-structures, in order to analyze and synthesize it, and to stimulate new reflections and studies on this innovative way of conceiving the organization of information and knowledge.

5. REFERENCES


