HP-Graph as a Basis of a DSM Platform Visual Model Editor

Nikolai M. Suvorov
Department of Business Informatics
National Research University
Higher School of Economics
Perm, Russian Federation
E-mail: SuvorovNM@gmail.com

Lyudmila N. Lyadova
Department of Business Informatics
National Research University
Higher School of Economics
Perm, Russian Federation
E-mail: LNLyadova@gmail.com

Abstract. The language-oriented approach is becoming more and more popular in the development of information systems, but the existing DSM-platforms that implement this paradigm have significant limitations, including insufficient expressive capabilities of the models used to implement visual model editors for complex subject areas and limited abilities to transform visual models. Visual languages are usually based on graph models, but the types of graphs used have restrictions, such as inefficiency and complexity of operations and insufficient expressiveness of the created models. For creating a tool that does not have the described constraints, development of a new formal model is needed. HP-graphs can become a solution for this problem, not only providing the ability to define and implement new visual languages, but also providing a basis for implementing operations on models built using these languages. Definitions of the HP-graph and its elements are given. Justification of expressive power of the HP-graph is presented. Main operations for the HP-graph are described. The chosen graph formalism unites expressive possibilities of various types of graphs and allows the creation of a flexible visual model editor based on it for a DSM-platform.

Keywords: Domain-specific language; DSM platform; visual model; graph model; HP-graph; algorithms on graphs.

I. INTRODUCTION

The study of any objects and processes, as well as their design, can barely be done without modeling, that is why software tools that allow non-IT specialists to build various models and formalize descriptions of objects and processes, or use modeling as a method of analysis for the study of objects are becoming more popular. Among the subject areas where modeling is particularly important, the development of information systems stands out. Currently the main approach to creating large information systems is a model-oriented approach [1]. Using it, developers usually deal only with models, which helps to ensure high quality of programs and prevent errors. CASE tools [2], which automate the system development process as much as possible due to the capabilities of visual modeling, model interpretation, and code generation based on the created models, are used especially for these purposes.

However, the traditional model-oriented approach to developing systems has its drawbacks, among which are:

- Universality of the languages used for system development as languages operate not in terms of the subject area, but in constructs of the means by which the system is created.
- Immutability of modeling languages, which does not allow all the subtleties, pitfalls, and limitations of the subject area to be displayed and taken into account.

- Complexity of modification of the created systems as making changes to the system is possible only if there are development tools, source codes and a professional IT specialist.
- The impossibility of transitioning from one modeling language to another, but, creating large systems usually involves building several models describing the system from different points of view, with different granularity, so in such cases, there is a need to harmonize the models created by different professionals at different stages of development, which requires the ability to perform a transition from one modeling language to another.

These problems are solved by a paradigm called language-oriented programming [3]. This paradigm at the initial stage of development implies the creation of a metamodel of a subject area represented by one or more languages for solving various project tasks. These languages are used to build the necessary models for implementing the system. For implementing this approach DSM-platforms [4], language tools, and Meta-CASE systems, such as MetaEdit+ [5], which facilitate the development of domain-specific languages (DSL), are usually used. These languages operate in terms of the subject area and reflect the specifics of the tasks they solve. Moreover, subject-oriented languages can also consider the qualifications of users who will use them [6].

Nevertheless, existing tools only partially solve the problems of the traditional approach to modeling. To solve all the problems described above, a language tool must meet the particular requirements [4], [7], [8]. It should

- have an ability to define modeling languages for most subject areas;
- have an ability to dynamically change the modeling language;
- have an ability to alienate the created modeling language from the system where it has been developed;
- have an ability to modify the visual model of the system, rather than the source code, when a modeled process or system undergoes changes;
- unify representation and description of both models and metamodels, which allows a person to work with models and metamodels using the same tools, as well as, for example, provides the opportunity to perform vertical and horizontal transformations of visual models.

To create a tool that has all these features, the development
of a new formal model is needed. Visual languages are usually based on graph models [1], but the types of graphs used have limitations, such as insufficient expressiveness of the created models, inefficiency, and complexity of operations. However, there is a more powerful formal model that solves these problems, but has not been used by developers yet, which is called a hypergraph with poles (HP-graph) [9], which connects the expressive capabilities of various types of graph models.

II. RELATED WORKS

Many different tools have been created that allow people to develop modeling languages and build models based on these languages. These tools are Microsoft DSL Tools [10], Eclipse Sirius [11], MetaEdit+ [5], Microsoft Visio [12], QReal [13], etc. Detailed description and comparison of these platforms is given in [4]. All these platforms have some limitations and do not fully meet the requirements described above. Consideration of main constraints of the platforms is given below.

Microsoft DSL Tools uses templates based on UML diagrams to create a new DSL, which leads to complexity and confusion when building model hierarchies and leads to appearance of limitations and inaccuracies in the resulting modeling language [8]. Moreover, this platform is characterized by the lack of the ability to dynamically change metamodels and transform models, as well as the inability to use DSL outside of MS Visual Studio.

Eclipse Sirius offers a solution for rapid development of a graphical tool for DSM, but certain complex tasks may require changes to the EMF and GMF code. There is also a need for interpreted expressions which will be evaluated at runtime to provide a behavior specific to domain and representations and which can only be written in Accelero, OCL or Java. Sirius allows a user to perform horizontal transformations, but the knowledge of special addons is needed.

MetaEdit+ contains only limited possibilities for transforming visual models. Models exported from the platform have their own format, which makes it difficult to use models created in this platform in other software tools.

The main drawbacks of Microsoft Visio are the inability to change the modeling language while the system is running, and the need to purchase MS Visio to use the tools developed on its basis. Also, a language metamodel can only be built using a UML class diagram, which significantly limits the platform's capabilities and complicates the process of creating languages.

The QReal platform does not have the ability to change the metalanguage, the ability to transform models, and this platform is characterized by the complexity of modifying the created modeling language.

As it seems from the Table I, there is no platform that meets all the previously given requirements. Nevertheless, it should be noted that at least some of the requirements for tools are met by each of the platforms listed. Modeling and designing information systems tend to be done using special methodological approaches which can be divided to structural and object-oriented approaches. Despite the difference in the approaches and the division of all tools into two large groups depending on the approach underlying them (UML and "No-UML"), there are general modeling principles that the model should be aimed to implement.

<table>
<thead>
<tr>
<th>Requirements</th>
<th>MS DSL Tools</th>
<th>Eclipse Sirius</th>
<th>MetaEdit+</th>
<th>MS Visio</th>
<th>QReal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to define modeling languages for most subject areas</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ability to dynamically change the modeling language</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ability to alterate the created language from the system</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ability to modify the visual model</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ability to perform a horizontal transformation</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

The essence of the structural approach is to decompose a process into automated functions – the function of the upper level is decomposed and divided into subfunctions, refining properties of the functions at the upper levels of the hierarchy. Each subfunction, in turn, is decomposed into elements of the next level, and this happens until the obtained structure becomes trivial enough. Among the diagrams of this approach are Structural Analysis and Design Technique (SADT), a Data-Flow Diagram (DFD) and an Entity-Relation Diagram (ERD). The structural approach is used in simulation systems [15], as well as for functional and information modeling [16].

The essence of the object-oriented approach is an object decomposition, when the system is represented as a set of objects that exchange messages during the interaction. Moreover, the object itself in this case is an independent entity characterized by its state, behavior, and semantics [18]. Based on this approach, a set of DSLs [19],[20] is developed, but this approach is characterized by certain disadvantages, among which the complexity of building a hierarchy of models is highlighted. Using this approach does not always allow a person to properly express the concepts of the subject area, so the resulting language may have some limitations and inaccuracies. However, using it we can significantly reduce the language development time [21]. With all this in mind, the formalism underlying the visual model editor for a DSM-platform must meet the following requirements:

- To allow multi-level and multi-aspect modeling, which makes the decomposition of models from different points of view possible.
- To unify the description of models at different levels of the hierarchy, which means that the same formalism should be used to describe both models and metamodels.
- To allow development of modeling languages for a wide range of subject areas.
- To allow a user to discard constructions that are not details of the subject area, which will simplify the study of the developed language by end users.
- To perform both horizontal and vertical transformations.

Various types of graph formalisms are used for constructing and visualizing models, including oriented
graphs, multigraphs [22], hypergraphs [23], hi-graphs [24], meta-graphs [7, 26], P-graphs [27, 28]. Nevertheless, all these formalisms cannot meet all the mentioned requirements due to their certain limitations, therefore, development of a new graph model is needed.

III. DESCRIPTION OF THE GRAPH MODEL

Hypergraph with poles (HP-graph) is a graph model which meets the given requirements and can be used as a base for a visual model editor.

HP-graph is an ordered triple \( G = (P, V, W) \), where

\[ P = \{p_1, \ldots, p_n\} \] is a set of external poles,

\[ V = \{v_1, \ldots, v_m\} \] is a non-empty set of vertices,

\[ W = \{w_1, \ldots, w_k\} \] is a set of edges [9].

Let \( Pol \) be an abstract set of all poles of the graph. Thus, \( 2^{Pol} \) is a powerset of all poles of the graph. Then:

- Every vertex \( v \in V \) is a subset of the set of all subsets of poles \( (v 
\in 2^{Pol}) \) but \( \forall \nu \in V, \forall \nu \in V \) \( w = v \cap \cap \nu \), which means that \( V \) is a set of mutually disjoint subsets of \( Pol \).

- A set of external poles \( P \) is also a subset of the powerset of poles \( (P 
\in 2^{Pol}) \). This set consists of input and output poles of the graph \( P = I(G) \cup O(G) \). Each vertex of the graph \( v \in V \) is also represented by a set of input \( (I(v)) \) and output \( (O(v)) \) poles \( \{p_{v_i}, \ldots, p_{v_n}\} \) \( (\forall v \in V) = \{p_{v_i}, \ldots, p_{v_n}\} \). Sets of input and output poles of any graph can be assigned. If no poles are specified for a vertex, it is assumed that the vertex consists of a single pole, which is both input and output \( (I(v) = O(v)) \).

- Each edge \( w \in W \) defines connections between vertices and is represented as a subset of the powerset of poles \( (w = P_e = \{p_{w_1}, \ldots, p_{w_n}\} \subseteq 2^{Pol}) \). An edge cannot be represented as an empty set \( (\forall w \in W : [P_e = \emptyset]) \). The edge can allow a vertex to be linked to itself. Each edge must contain at least one input pole and one output pole, so for \( \forall v \in V(G) \), \( \forall w \in W(G) \) the following condition must be met: \( [\exists p \in w \cap I(v) \cap O(G)] \) and \( [\exists r \in w \cap I(v) \cap O(G)] \).

An example of the hypergraph with poles is demonstrated on Fig. 1.

![Fig. 1. Example of an HP-graph](image)

On this figure external poles are represented by a set \( P = \{p_1, \ldots, p_s\} \), edges are represented by a set \( W = \{w_1, \ldots, w_k\} \) and vertices are represented by a set \( V = \{v_1, v_2, v_3\} \). Every vertex contains a certain number of poles \( p \), which are connected to other poles by means of hyperedges from the set \( W \).

In the HP-graph, edges and vertices are represented as sets of inputs and outputs, while the actual structure of these elements is hidden. Thus, it can be assumed that these elements are represented as a "Black box".

A. Main Operations

To describe main operations on the HP-graph, let us define \( G = (P, V, W) \) as an original HP-graph, \( G = (P, V, W) \) as a resulting HP-graph, \( v \) as a vertex, \( p_1 \) as an inner pole, \( p_2 \) as an outer pole and \( w \) as an edge.

The following operations add elements to an HP-graph:

- \( v + p_1 \) is the addition of the inner pole to the node. The pole is added to both the vertex itself and the set of all poles of the graph:

\[ Pol(G) = Pol(G) \cup \{p_1\} \]

- \( G + v \) is the addition of the node to the graph. If a cardinality of \( v \) is more than 0, a vertex is added to the set of vertices \( V(G) \), and all poles of this vertex are added to the set of all poles of the graph:

\[ Pol(G) = Pol(G) \cup v \]

- \( W \) is the addition of the edge to the graph. An edge is formed from the already existing poles of the graph by combining them into a single set. Let \( w \) be the set of input poles of an edge \( w \), and \( O(w) \) be the set of output poles of \( w \), then the operation is represented as:

\[ W(G) = \{W(G) \cup \{w\} \mid [O(w) \neq \emptyset] \}

- \( w + p_1 \) is the addition of the inner pole to the edge. An existing pole belonging to one of the vertexes is added to the edge \( (\forall v \in V(G) \mid p_1 \in v) \), which is represented as:

\[ w' = w \cup \{p_1\} \]

- \( w + p_2 \) is the addition of the outer pole to the edge. An existing pole belonging to the set of outer poles \( (p_2 \in P(G)) \) of the graph is added to the edge:

\[ w' = w \cup \{p_2\} \]

- \( G + p_2 \) is the addition of the outer pole to the graph. A pole is added to both the set of outer poles of the graph \( G \) and the set of all poles:

\[ Pol(G) = Pol(G) \cup \{p_2\} \]

\[ P(G) = P(G) \cup \{p_2\} \]

The following operations remove elements from an HP-graph:

- \( v - p_1 \) is the removal of the inner pole from the node. When a pole is removed from a vertex, all its occurrences in the edges are cut off and it is removed from the set of all poles of the graph:

\[ \forall w \in W(G) \mid \{w \mid \{p_1\} \} \]

\[ Pol(G) = Pol(G) \setminus \{p_1\} \]

\[ v' = \{v \setminus \{p_1\} \mid \{p_1\} \} \]
• $G - v$ is the removal of the node from the graph. In addition to deleting a vertex, all occurrences of the poles of this vertex in the edges are cut off, and all poles of the vertex are removed from the set of poles of the graph:

$$\forall w \in W(G) \forall p \in v \{ w = \{ w \setminus \{ p \} \} \},$$

$$Pol(G') = Pol(G) \setminus \{ v \},$$

$$V(G') = V(G) \setminus \{ v \}.$$  

• $G - w$ is the removal of the edge from the graph. Performing this operation only removes an edge from the set of all edges of the graph, leaving the external poles and vertex poles unchanged:

$$W(G) = W(G) \setminus \{ w \}.$$  

• $w - p1$ is the removal of the inner pole from the edge. A pole is removed only from an edge $w$, without changing the set of all poles of the graph and the set of poles of the vertex to which it belongs, but if the resulting edge $w'$ does not contain at least one input and one output pole, then the edge is removed:

$$w' = w \setminus \{ p1 \},$$

$$W(G) = \{ W(G) \setminus \{ w \} \} \setminus \{ I(w) = 0 \text{ or } O(w) = 0 \}.$$  

• $w - p2$ is the removal of the external pole from the edge, which is equal to the previous operation.

• $G - p2$ is the removal of the outer pole from the graph, which also includes removing this pole from all edges that contain this pole:

$$\forall w \in W(G) \{ w = \{ w \setminus \{ p2 \} \} \},$$

$$P(G) = P(G) \setminus \{ p2 \},$$

$$Pol(G) = Pol(G) \setminus \{ p2 \}.$$  

B. Operations of Decomposition

A hypergraph with poles allows vertices and edges to be decomposed during the decryption operation. This feature makes multilevel representation possible. This possibility is achieved by correctly correlating the poles of the source and received graph which is done by implementing a mapping function.

The mapping $f: v \rightarrow P$, which is a decoding function for a vertex $v$, must be concordant with the sets $I(v)$ and $O(v)$, so that $\forall p \in I(v): \{ f(p) \in I(G) \}$, $\forall r \in O(v): \{ f(r) \in O(G) \}$. Thus, the mapping of the pole $p \in v$ to the next level of the hierarchy is represented as $f(p) = \pi$, where $\pi \in P(G)$, which means that a pole $p$ becomes the external pole $\pi$ for a resulting graph.

Fig. 2 illustrates decomposition of the vertex $v3$ by a new HP-graph.

An edge can be decomposed similarly but with the help of mapping functions $f: w \rightarrow P$ which also must be concordant with sets of input ($I(w)$) and output ($O(w)$) poles, so that $\forall p \in I(w): \{ f(p) \in I(G) \}$, $\forall r \in O(w): \{ f(r) \in O(G) \}$. Thus, the mapping of the pole $p \in w$ to the next level of the hierarchy is also represented as $f(p) = \pi$, where $\pi \in P(G)$. Example of decomposition of the edge $w6$ is demonstrated on Fig. 3.

As is seen, the decomposition of edges and vertices is almost equal; therefore, it is possible to define a common decryption algorithm for these structures.

To do it, let us define a set of structures $Str = V \cup W$. Hence, $str \in Str$ is a structure which can be either a vertex or an edge. The algorithm of the structure decomposition by a new HP-graph can be described as follows:

**Algorithm 1. Procedure DecomposeStructure**

```plaintext
G = new HPGraph();
foreach p \in Str:
    if (p \in I(str)):
        I(G) = I(G) \cup p;
    if (p \in O(str)):
        O(G) = O(G) \cup p;
    Open_{str} = Open_{str} \cup (str, G)
```

As is seen from the algorithm, for every structure $str$ several decoding operations can be defined. Generally, they can be presented as $Open_{str} \subseteq str \times G_{all}$, where $G_{all}$ is the set of all HP-graphs determined on the set $Pol$.

An edge of an HP-graph can also be decrypted by ordinary links between the poles. To implement this operation, it is necessary to define the set $E_w = \{ e_1, \ldots, e_n \} \subset I(w) \times O(w)$ for each edge $w \in W$, so that every link $(e \in E_w)$ is represented by a pair $(p, r)$ provided that $p \in I(w)$, $r \in O(w)$. Thus, the decoding of the edge $w \in W$ can be represented by the mapping function $f: w \rightarrow E_w$, which replaces the hyperedge with normal connections between the input and output poles.

Fig. 4 illustrates the example of hyperedge decoding by ordinary links. As $E_w \subseteq I(w) \times O(w)$, some input and output poles can be unconnected such as poles $p9[O]$ and $p11[I]$ in Fig. 4.
C. Operations of Transformation

Many different approaches are used to transform visual models, but from the point of view of some scientists and developers [8], [29] the most promising one is the algebraic approach [30], which allows parsing graphs and checking graph models for consistency. This approach and its modifications are implemented in such tools as MetaLanguage [8], AOG [31] and VIATRA [32]. It is worth mentioning that there are also toolsets, such as ATL [14], that implement technologies from other areas of software engineering, but most of them have considerable restrictions.

To define transformation operations, it is necessary to give a definition to a subgraph of HP-graph. A subgraph of the HP graph \( G = (P, V, W) \) is an HP-graph \( G' = (P', V', W') \) that is part of the graph \( G (P' \subseteq P \ & \ \forall v \in P' \ \exists v \in V: \ [v' \in V] \ & \ W' \subseteq W) \) and fulfills the condition \( \text{Open}' \subseteq \text{Open} \). The subgraph must also meet the condition (1) to make transformation operations possible.

\[
(\forall v \in V \setminus V_{\text{partial}} \ [p \in v] \ \& \ \exists w \in W [p \in w]) \rightarrow w \in W'
\] (1)

The set \( V_{\text{partial}} \) is a set of the incomplete vertices in the graph, where \( V_{\text{partial}} \subseteq V \).

A subgraph can contain vertices called incomplete whose sets of poles can only be part of the sets of poles of the vertices of the original graph.

To perform a transformation, it is needed to select the source and the target graph and set production rules that describe the transformation. A production rule is represented as \( p = (G_l, G_k) \), where \( G_l \) is a pattern-graph and \( G_k \) is a replacement graph. Nevertheless, there is a restriction which is represented in (1) and must be satisfied to perform a transformation. It can be explained by the fact that it is unknown how certain hyperedges should change during the transformation while this restriction obliges to redefine all edges which are incident to poles involved in the transformation.

To display all hyperedges, all the poles that are included in them must be displayed, so it is needed to add such auxiliary (incomplete) vertices that store only those poles that belong to the displayed hyperedges.

An algorithm for the transformation can be divided into two functions. The first one removes a subgraph isomorphic to the pattern and the second one adds replacement graph to the original graph.

The first step can be described as follows:

```
Algorithm 2. Function DeleteGraph(HostG, G_k)
G' = Find_Isomorphic_Subgraph(HostG, G_k);
if partials = {};
foreach w \in W(HostG):
    W(HostG) = W(HostG) \ {w'};
foreach v \in V(HostG):
    if (v \in V(HostG));
        V(HostG) = V(HostG) \ {v'};
else:
    partials = partials \ {v'};
foreach p' \in P'(HostG):
    if (w' \in W(HostG)[p'] \in w):
        P'(HostG) = P'(HostG) \ {p'};

return partials;
```

The second step of the algorithm will be following:

```
Algorithm 3. Procedure AddGraph(HostG, G_k, partials)
foreach p \in P(HostG):
    if p \in P(HostG):
        P(HostG) = P(HostG) \ {p};
foreach v \in V(HostG):
    if (v \in V(HostG));
        V(HostG) = V(HostG) \ {v};
foreach w \in W(HostG):
    W(HostG) = W(HostG) \ {w};
```

These algorithms can be repeated several times as the set of transformation rules may not be limited to just one rule.

IV. JUSTIFICATION OF EXPRESSIVE POWER OF FORMALISM

It is possible to justify the transcending expressive power of the HP-graph by proving that the graph formalisms generally used for building and visualizing models can be represented as an HP-graph. Previously, it was mentioned that oriented graphs, hypergraphs, hi-graphs, meta-graphs and P-graphs are most frequently used for such purposes. Table II describes formulas which represent any of these graph structures as an HP-graph.

```
TABLE II. REPRESENTATION OF GRAPHS AS AN HP-GRAPH

<table>
<thead>
<tr>
<th>Graph model</th>
<th>Representation in the HP-graph G' = (P', V', W')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directed Graph</td>
<td>G = (V, E), E = W, E \ {v' \in V}</td>
</tr>
<tr>
<td>Hypergraph</td>
<td>G = (X, E), E = W, \forall v' \in V:</td>
</tr>
<tr>
<td>Hi-graph</td>
<td>G = (X, E), E = W, \forall v' \in V:</td>
</tr>
<tr>
<td>Metagraph</td>
<td>G = (V, MV, E), E = W, MV = W</td>
</tr>
<tr>
<td>P-graph</td>
<td>G = (P, V, W), P = P, V = V, W = W, \forall w' \in W:</td>
</tr>
</tbody>
</table>

From the table it can be concluded that the HP-graph has more expressive power than the previously described graph models. These graph models are special cases of the HP-graph; thus, the HP-graph is a generalization of all of these graph formalisms.

V. CONCLUSION

The definition of the mathematical apparatus underlying the visual model editor was given above, including a detailed description of the graph structure itself, as well as the
operations that can be performed on it. For the selected graph formalism, algorithms for decoding vertices and edges, as well as algorithms for performing transformations, were described.

The HP-graph unites expressive possibilities of various types of graphs, therefore, algorithms that are designed for these types of graphs (particularly model transformation algorithms [33], [34]) can also be implemented for HP graphs. The time complexity of model transformation algorithms can be reduced. The paper proves that HP-graph allows the creation of a flexible visual model editor based on this graph formalism for a DSM-platform. Representing both vertices and links as sets of poles simplifies the object model of DSM editor and visual model editing algorithms.

It is planned to develop a program that will demonstrate the practical significance of the selected graph formalism.

REFERENCES