Abstract — Usage of visual domain-specific languages in software engineering allows to simplify the process of software creation and to attract to it the experts in domain, who are not professional programmers. However creation new domain-specific language is the nontrivial task, therefore the problem of automation of their development process is the topical task. For the automation, designing of visual modeling languages it is offered to use the ontologies received as a result of the analysis of text corpus. In article, the approach to automatic creation of visual modeling languages on the basis of domain ontologies is considered.

Keywords — domain-specific modeling languages; ontologies; DSM-platform; MetaLanguage; metagraphs.

I. INTRODUCTION

The domain-specific modeling languages (DSMLs), which are designed to solve a particular class of problems in a specific domain, are increasingly used at software development and maintenance process. Unlike the general-purpose languages, DSMLs are more expressive, easy to use and clear to various categories of users, as they operate with domain terms, which are familiar to users [1, 2]. For this reason now a large number of DSMLs is designed for creation of systems in different domains: artificial intelligence systems, distributed systems, mobile applications, real-time and embedded systems, simulation systems, etc. [3-7].

Despite of all DSMLs advantages they have one big disadvantage – complexity of their designing. If general-purpose languages allow to create models irrespectively to domain, in case of DSMLs for each domain, and in some cases, for each task, it is necessary to create a new domain-specific language. Another shortcoming of visual domain-specific language is that it is necessary to create convenient graphical editors to work with it. Therefore, a problem of automation of DSMLs development process is rather topical.

To support the process of development and maintenance of DSMLs the special kind of software — language workbench (DSM-platform) is used. Usage at DSMLs creation of a language workbench considerably simplifies the process of their designing. There are various DSM-platforms for creating visual DSMLs with the ability of determining user’s graphical notation: MetaEdit+, Microsoft DSL Tools, Eclipse GMF, QReal, etc. However, these tools do not allow to automatic create DSMLs.

This problem can be solved by the development of methods and tools, which will allow on the basis of a set of documents, available in domain, to build a conceptual domain model and automatically design a modeling language, which corresponds to singularities of domain, needs of various categories of users [8].

Let’s consider the most advanced language workbenches [9].

II. RELATED WORKS

MetaEdit+ is a multiplatform language workbench that enables users to simultaneously work with several projects, each of which can have a few models [10]. At usage of this DSM-platform, besides a possibility of domain-specific language creation, the developer receives the CASE tool into which this language is integrated. MetaEdit+ allows to use several DSMLs at system creation.

The approach based on metamodels (models of modeling languages) interpretation, instead of code generation, which is used in MetaEdit+ allows changing the DSMLs definition at run-time.

DSL Tools [11] and Eclipse GMF [12] technologies provide the user with advanced IDE MS Visual Studio and Eclipse respectively. State Machine Designer [13], in fact, is an add-on to DSL Tools, which eliminates some of its shortcomings. However, the State Machine Designer allows creating DSMLs only using UML Activity Diagrams that considerably limits the range of solving tasks.

Multiplatform system QReal [14, 15] allows to define metamodels both in visual and textual view, therefore developers have a possibility to select the most suitable for them format of language description representation. Availability of an interpreter of behavioral diagrams and a debugger of the generated code puts this system in one row with tools MS DSL Tools, Eclipse GMF, which use for these purposes IDE.
The analysis of DSM-platforms has revealed the following
main restrictions inherent in the majority of the considered
systems [16]:

1. Impossibility of multilevel modeling. Presence of such
   possibility would allow making changes at metalanguage
description, to extend it with new
   constructions, thus bringing the metalanguage to the
   specifics of domain.

2. Modification of DSMLs description leads to necessity
   of regeneration of language editor: for modification
   DSMLs at first it is necessary to change its metamodel,
to regenerate the source code of the editor, and only
   then it is possible to begin build models.

3. “Excess” functionality of the language workbench,
   which is not used at DSMLs creation. This functionality
   complicates the study of tools by the users, which are
   not professional programmers.

4. Lack of tools of horizontal models transformation.
   These means allow not only to create unified system
description on the basis of the models constructed at
   various stages of system development, but also to
generate source code according to user-specified
template or to make conversion of the model described
with one modeling language to model fulfilled in other
graphical notation.

Moreover, the considered tools do not allow to fulfill
automatic construction of DSMLs on the basis of domain
analysis. This possibility allows to:

- simplify the process of DSMLs creating;
- create DSMLs, approached to specific domain;
- attract to DSMLs development process the users, who
  are not the professional programmers.

Thus, it is possible to say that creation of methods and tools
of automatic DSMLs creation is the topical task, the solution of
which will significantly simplify the process of visual domain-
specific languages designing for various domains, and also will
submit a possibility to involvement the experts in the process of
development and maintenance of information systems.

For development of methods and tools for automatic
visual DSMLs designing, it is necessary to solve the following
tasks:

1. Construct a mathematical model, which will allow to
   unify describe domain ontologies and metamodels of
   visual languages.

2. Develop rules of ontologies transformation in
   constructions of visual modeling languages.

3. Develop metrics and methods of comparing of
   DSMLs, which submit a possibility to estimate
   proximity of generated automatically domain-specific
   languages to domain specificity.

4. Implement the developed methods in dynamic library.

5. Integrate the created library into the MetaLanguage
   system.

6. Perform approbation of the received results by
development of visual modeling languages for various
domains.

III. METALANGUAGE SYSTEM

The MetaLanguage system eliminates some restrictions of
the considered DSM-platforms.

This language workbench is designed to create visual
dynamic adaptable domain-specific modeling languages,
construct models using these languages and transform of the
created models in various textual and graphical notations
[17, 18].

A. Metalanguage of MetaLanguage System

One of the basic elements of language workbench is the
metalanguage (meta-metamodel) – language for describing of
other languages (metamodels). Thanks to presence of
metalanguage, the DSM-platform allows to create domain-
specific languages for the various domains that operate with
familiar to user concepts. The main difference between
metamodels of MetaLanguage system from the MOF (Meta
Object Facility) approach, used in the majority of DSM-
platforms, is that thanks to interpretation of models at various
abstraction levels, instead of the source code generation on
their basis, it is possible to modify of DSML’s constructions in
dynamic, during models creation. Besides, the process of
metamodel creation becomes multilevel, so having defined a
metamodel and having selected it as a metalanguage, the
developer can use this meta-metamodel for creation of other
metamodels, and this process can be infinite.

The basic elements of the metalanguage of MetaLanguage
system are entity, relationship and constraint [19].

The entity describes a particular construction of modeling
language, i.e. it is the domain object, important from the point
of view of the solving problem.

Visual language constructions in rare cases exist
independently, more often they are in some way related to each
other, therefore at metamodel creation importantly not only to
define the basic language constructions, but also correctly
specify the relationships. The relationship is used for
describing a physical or conceptual links between entities.
Metamodel allows to create three types of relationships:
association, aggregation, inheritance.

In practice quite often, there are cases when it is necessary
 to impose some constraints on entities and relationships
 between them. Some of constraints are set by metamodel
structure, and others are described on some language.

Let’s consider an example. Fig. 1 shows a fragment of
metamodel for UML Use Case diagrams. The metamodel
contains two entities “Actor” and “Use Case”.

The entity “Use Case” has following attributes: “Name”,
“Description”, “Creation_Date”. The attribute “Name” has a
string type and defines the Use Case name. The attribute
“Description” sets the short description of the Use Case. The attribute of entity “Actor” is a string attribute “Name”, which specifies the name of the Actor.

![Use Case Diagram](image)

Fig. 1. Fragment of metamodel for UML Use Case diagrams

**B. Architecture of MetaLanguage System**

The architecture of MetaLanguage system is presented in Fig. 2. Uniform storage of all information about the system is the repository. It contains information about metamodels, models, entities, relationships, attributes, constraints. Information about the models and metamodels is stored uniformly, that allows to work with it by a single tool. The browser of models allows to load/save metamodels together with the models, created on their basis, to fulfill over metamodels and models various operations (editing, constraint checking, transformation, etc.). The graphical editor is the component, which provides the user the tools for metamodels and models creation. The validator allows to check constraints specified by user at metamodel describing. The transformer is the component that provides the ability to fulfill horizontal transformations of models to text on target programming language or to visual models, described in other graphical notation.

![Architecture Diagram](image)

Fig. 2. Architecture of MetaLanguage system

Having described the basic components of a MetaLanguage system, let’s consider how visual domain-specific modeling languages are designed.

Process of DSML definition begins with metamodel creation. For this purpose, it is necessary to specify the main constructions of created language, to define relationships between them, to set constraints imposed on the metamodel entities and relationships. After building of metamodel the developer gets a customizable extensible visual modeling language.

Then using created DSML, the user should design models containing objects that describe specific domain concepts and links between them.

The validator should check up whether model satisfies to constraints, which were imposed on metamodel elements.

Then the developer can save the constructed metamodels and models in the form of XML-files or transform these models to other textual or graphical notation.

At metamodel modification, the system automatically makes all necessary changes in the models, which are created on the basis of this metamodel.

For automatic creation of DSMLs, it is necessary that the language workbench on the basis of domain description fulfilled creation of language’s metamodel. Since the structure of the metamodel of MetaLanguage system is similar to ontologies description languages, then it was decided to use ontologies as the basis of automatic DSMLs designing. There are various systems for the automatic construction of ontologies on the basis of a text corpus: OwlExporter [20], OntoGrid [21], etc. These systems allow to fulfill ontology creation on the basis of initial set of documents. The resulting ontology will be used by the MetaLanguage system for automatic creation of visual DSMLs (see Fig. 3).

**IV. AUTOMATIC CREATION OF DSMLS**

Formally ontology is the tuple $O = \{T, R, I\}$, where

- $T$ is the finite nonempty set of domain concepts;
- $R$ is the finite set of relationships between concepts;
- $I$ is the finite set of interpretation of ontology concepts and relationships.

Thus, it is possible to say that the ontology is a directed labeled graph. The metamodel of visual modeling language is a graph also. For this reason the formalism that allows to describe domain ontology and metamodel of visual DSML in unified form is a labeled graph.

There are several types of graphs that can be used for representation of visual languages and ontologies: the classical graphs, digraphs, multigraphs, pseudographs, hi-graphs, hypergraphs, metagraphs and others.

![Ontology Diagram](image)

Fig. 3. Automatic creation of metamodel of DSML
As an analysis result of various types of graph it has been defined that the most appropriate formalism for describing the syntax of visual modeling languages in MetaLanguage system are pseudo-metagraphs [22].

Metagraph is an ordered pair $G = (V, E)$, where $V$ is a finite nonempty set of nodes, $E$ is a set of edges. Each edge $e_i = (V_i, V_j), V_i, V_j \subseteq V$ connects two subsets of nodes.

This type of graphs allows to reduce the number of graph arcs and to make the model more structured, logical. Therefore, in the metamodel graph the attributes and constraints of each entity and relationship are united in the separate sets. The model becomes more demonstrative, clear and corresponding to the solving task.

### A. Metamodel Graph

Let’s describe with usage of this formalism a metamodel of a visual modeling language.

Let $Ent = \{ent_i\}, i \in \mathbb{N}$ ($\mathbb{N}$ is a set of natural numbers) is a set of metamodel entities, number of set elements is potentially unlimited, but at every fixed point in time is finite.

The set of metamodel relationships denotes as $Rel = \{rel_k\}, k \in \mathbb{N}$, number of set elements is potentially unlimited, but at every fixed point in time is finite.

Let’s introduce the following designations:

- $EAttr = \{eattr_{i\ j}\}, i = 1, |Ent|, j \in \mathbb{N}$ is the set of metamodel graph nodes, which corresponds to entities attributes;
- $RAttr = \{rattr_{i\ k}\}, k = 1, |Rel|, l \in \mathbb{N}$ is the set of metamodel graph nodes, which corresponds to relationships attributes;
- $ERest = \{erest_{i\ j}\}, i = 1, |Ent|, j \in \mathbb{N}$ is the set of metamodel graph nodes, which corresponds to constraints imposed on entities;
- $RRest = \{rest_{i\ k}\}, k = 1, |Rel|, l \in \mathbb{N}$ is the set of metamodel graph nodes, which corresponds to constraints imposed on relationships;
- $EEA = \{eea_i\}, i = 1, |Ent|$ is the set of metamodel graph arcs connecting each entity with the set of its attributes;
- $EERA = \{era_{i\ j}\}, k = 1, |Rel|$ is the set of metamodel graph arcs connecting each relationship with the set of its attributes;
- $EER = \{eer_{i\ j}\}, i = 1, |Ent|$ is the set of metamodel graph arcs connecting each entity with the set of its constraints;
- $ERR = \{err_{i\ j}\}, k = 1, |Rel|$ is the set of metamodel graph arcs connecting each relationship with the set of its constraints;
- $EERR = \{eerr_i\}, i \in \mathbb{N}$ is the set of arcs corresponding to links between entities and relationships.

The number of elements of sets $EAttr$, $RAttr$, $ERest$, $RRest$, $EEA$, $ERA$, $EER$, $ERR$, $EERR$ potentially is not limited, but it is finite at every fixed point in time.

The metamodel graph is the directed pseudo-metagraph $GMM = (V, E)$, where $V$ is a nonempty set of graph nodes, $E$ is set of graph arcs and these sets are defined by (1) and (2):

$$V = \bigcup_{i=1}^{|Ent|} EAttr \cup \bigcup_{i=1}^{|Ent|} ERR \cup \bigcup_{i=1}^{|Rel|} ERest$$

$$\bigcup_{i=1}^{|Rel|} RAttr \cup \bigcup_{i=1}^{|Rel|} RRest_i$$

(1)

$$E = EEA \cup EERR \cup ERA \cup ERR \cup EER$$

(2)

Let’s consider an example. We will construct a metamodel graph for the entity “Use Case” of UML Use Case diagrams. Metamodel of this diagram type is shown in Fig. 1. Attributes of the entity “Use Case” are “Name”, “Description”, “Creation_Date”, i.e. for given entity $EAttr = \{“Name”, “Description”, “Creation_Date”\}$.

The metamodel graph corresponding to a fragment of the “Use Case” entity is shown in Fig. 4.

![Fig. 4. Fragment of metamodel graph for “Use Case” entity](image)

As can be seen from the figure $ERest_i = \emptyset$, $EEA = \{eea_i\}$, $EER = \emptyset$, $ERR = \emptyset$.

### B. Ontology Graph

Let’s introduce the following designations:

- $T = \{t_i\}, i \in \mathbb{N}$ is the set of ontology graph nodes, which corresponds to concepts of ontology;
- $OAttr = \{oattr_{i\ j}\}, i = 1, |T|, j \in \mathbb{N}$ is the set of ontology graph nodes, which corresponds to concept attributes, which are not relationships;
- $OInst = \{oinst_{i\ j}\}, i = 1, |T|, j \in \mathbb{N}$ is the set of ontology graph nodes, which corresponds to concept instances;
The number of elements of sets \( T, OAttr, OInst, ORel, OA, OI, TR \) potentially is not limited, but it is finite at every fixed point in time.

The ontology graph is the directed pseudo-metagraph \( GO = (VO, EO) \), where \( VO \) is a nonempty set of graph nodes, \( EO \) is set of graph arcs and these sets are defined by (3) and (4):

\[
VO = T \cup \left( \bigcup_{i=1}^{n} (OAttr \cup OInst) \right) \cup ORel,
\]
\[
EO = OA \cup OI \cup TR.
\]

C. Mapping of Ontology Graph to Metamodel Graph

At mapping of a domain ontology to a modelling language metamodel it is necessary to fulfill the following algorithm:

1. To eliminate synonymy (to merge nodes containing synonymic concepts).
2. To delete instances of concepts and “is instance” relationships.
3. For each concept of ontology to create in a metamodel an entity with the concept’s attributes.
4. For each “is a” relationship of ontology to create an inheritance relationship in a metamodel.
5. For each “is part of” relationship of ontology to create an aggregate relationship in a metamodel.
6. For other relationships of ontology to create an association relationship in a metamodel.

According to this algorithm, let’s determine the mapping of the ontology graph to the metamodel graph, this mapping corresponds to operation of automatic creation of the metamodel graph.

Let’s introduce the following designations:

- \( ORel = \{orel\}, i \in \mathbb{N} \) is the set of ontology graph nodes corresponding to relationships between concepts.
- \( OA = \{oa_i\}, i = 1, |T| \) is the set of ontology graph arcs connecting each concept with the set of its attributes.
- \( OI = \{oi_i\}, i = 1, |T| \) is the set of ontology graph arcs connecting each concept with the set of its instances.
- \( TR = \{tr_i\}, i \in \mathbb{N} \) is the set of ontology graph arcs corresponding to links between concepts and relationships.
- \( OInstR = \{oinstr\}, i \in \mathbb{N} \) is the set of ontology graph nodes corresponding to “is instance” relationships between concepts.
- \( RInh = \{rinh\}, i \in \mathbb{N} \) is the set of metamodel graph nodes corresponding to inheritance relationships.
- \( RAggr = \{raggr\}, i \in \mathbb{N} \) is the set of metamodel graph nodes corresponding to aggregation relationships.
- \( RAssoc = \{rassoc\}, i \in \mathbb{N} \) is the set of metamodel graph nodes corresponding to association relationships between concepts.

The number of elements of sets \( OInh, OAggr, OInstR, RInh, RAggr, RAssoc \) potentially is not limited, but it is finite at every fixed point in time.

The mapping \( conc: T \rightarrow Ent \) for each ontology concept puts in correspondence a metamodel entity.

The mapping \( attr: OAttr \rightarrow EAttr \) for each ontology concept’s attribute puts in correspondence a metamodel entity’s attribute.

The mapping \( inh: OInh \rightarrow RInh \) for each ontology “is a” relationship puts in correspondence a metamodel inheritance relationship.

The mapping \( aggr: OAggr \rightarrow RAggr \) for each ontology “is part of” relationship puts in correspondence a metamodel aggregation relationship.

The mapping \( assoc: ORel / (OInh \cup OAggr \cup OInstR) \rightarrow RAssoc \) for each other ontology relationship puts in correspondence a metamodel association relationship.

The mapping \( ar: OA \rightarrow EEA \) for each arc \( oa_i \) of ontology graph puts in correspondence an arc \( eea_i \) of metamodel graph.

The mapping \( rr: TR \rightarrow ERR \) for each arc \( tr_i \) of ontology graph puts in correspondence an arc \( err_i \) of metamodel graph.

Thus, the creation of the metamodel graph on the basis of ontology graph is determined by mappings \( conc, attr, inh, aggr, assoc, ar, rr \).

D. “Smart House” Description Language

Let’s consider an example. Suppose that it is necessary to construct a visual modeling language for creation of models of “Smart House” systems.

At first, let’s analyze the components, which can be a part of “Smart House” systems. The basic elements of systems of this type are:

- life-support systems: heating, air conditioning and ventilation, lighting, security;
- sensors (devices that are responsible for obtaining of various readings and their sending to a central panel):
motion, leakings, fire and a smoke, closing/opening of object;
- system management tools: voice control, remote control (from a remote computer, phone, etc.), touch control (control by using of the touch screen of a central panel);
- central panel, which is responsible for receiving of data from sensors, management of life-support systems and obtaining of commands from the user.

The ontology received as a result of the analysis of a text corpus is presented in Fig. 5.

Fig. 5. Ontology of the structure of “Smart House” system

Concept “Life-support system” is the parent for concepts “Heating system”, “Air conditioning and ventilation system”, “Lighting system”, “Security system”. Concept “Sensor” is the parent for concepts corresponding to all types of system sensors. The relationship “Send information” connects the concept “Sensor” with the concept “Central panel”. The relationship “Interact” describes the interaction of the concepts “Life-support system” and “Central panel”. The relationship “Fulfill control” connects the concepts “System management tool” and “Central panel”.

In this case, according to the considered earlier algorithm, the MetaLanguage system has constructed the metamodel presented in Fig. 6.

Fig. 6. Metamodel of “Smart House” Description Language

Fig. 7 shows one of many possible models of “Smart House” system, constructed in MetaLanguage system with the usage of designed DSML.

Fig. 7. Model of “Smart House” system
V. CONCLUSION AND FUTURE WORKS

The DSMLs creation is a difficult task. To automate the designing of visual modeling languages the approach, allowing on the basis of domain ontology to generate a language description, is offered. It allows reducing complexity and time of DSMLs development, and gives the chance to nonprofessional programmers to develop their own languages. In the future, it is planned to develop metrics and methods of comparing of DSMLs, which submit a possibility to estimate proximity of generated automatically domain-specific languages to domain specificity and to generate visual DSMLs for other domains.

REFERENCES