VisualLISA: A Domain Specific Visual Language for Attribute Grammars

Nuno Oliveira¹, Maria João Varanda Pereira², Pedro Rangel Henriques¹, Daniela da Cruz¹, and Bastian Cramer³

¹ University of Minho - Department of Computer Science, Campus de Gualtar, 4715-057, Braga, Portugal, {nunooliveira,prh,danieladacruz}@di.uminho.pt
² Polytechnic Institute of Bragança, Campus de Sta. Apolónia, Apartado 134 - 5301-857, Bragança, Portugal, mjoao@ipb.pt
³ University of Paderborn - Department of Informatics, Fürstenallee 11, 33102, Paderborn, Germany, bcramer@upb.de

Abstract. The focus of this paper is on the translation of AG formalisms into a new visual language, and on the development of the associated programming environment. We present a solution for rapid development of VisualLISA editor using DEVIL. This tool uses traditional attribute grammars, to specify the language's syntax and semantics, extended by visual representations to be associated with grammar symbols. From these specifications a visual programming environment is automatically generated. This environment allows us to edit a visual description of an AG that is automatically translated into textual notations. In this paper, we emphasize the design and implementation of VisualLISA that is intended to be helpful for beginners and rapid development of small AGs.

1 Introduction

Attribute Grammars (AGs) [1] introduced by Knuth, 40 years ago, is a powerful and well-known formalism used to create language processors. An AG can be formally defined as the following tuple: \( AG = (G, A, R, C) \), where \( G \) is a context-free grammar, \( A \) is the set of attributes, \( R \) is the set of evaluation rules, and \( C \) is the set of contextual conditions. Each attribute has a type, and represents a specific property of a symbol \( X \); we write \( X.a \) to indicate that attribute \( a \) is an element of the set of attributes of \( X \), denoted by \( A(X) \). For each \( X \) (terminal or non-terminal), \( A(X) \) is divided into two disjoint sets: the inherited and the synthesized attributes. Each \( R \) is a set of formulas, like \( X.a = \text{func}(..., Y.b, ...) \), that define how to compute, in the precise context of a production, the value of each attribute. Each \( C \) is a set of predicates, \( \text{pred}(..., X.a, ...) \), describing the requirements that must be satisfied in the precise context of a production.

As can be deduced from this complex definition of AGs they are not as easy to specify as people would desire because there is a gap between the problem solution (the desired output) and the source language that must be interpreted. The user must take
care on choosing the appropriate attributes and their evaluation rules. Since the beginning, the literature related with compilers presents AGs using syntax trees decorated with attributes. So it is usual to sketch up on paper trees with attributes representing an AG. This strategy allows the developers to imagine a global solution of the problem (in a higher abstraction level) and to detect complex dependencies between attributes, symbols and functions, avoiding spending time with syntax details. However, such informal drawings require the designer to translate them manually into the input notation of a compiler generator. The person who drew it must go through the translation of the pencil strokes into the concrete syntax of the compiler generator. These inconveniences make the developers avoid the usage of AGs and go through non systematic ways to implement the languages and supporting tools. So, in this paper, we develop a Visual Language (VL), as a meta-language to write AGs, based on a previous conceptualization that we have proposed in [2].

VLs and consequently the Visual Programming Languages (VPLs) aim at offering the possibility to solve complex problems by describing their properties or their behavior through graphical/iconic definitions [3]. Icons are used to be composed in a space with two or more dimensions, defining sentences that are formally accepted by parsers, where shape, color and relative position of the icons are relevant issues. A visual programming language implies the existence of a Visual Programming Environment (VPE) [4, 5], because its absence makes the language useless. Commonly, a visual programming environment consists of an editor, enriched by several tools to analyze, to process and to transform the drawings.

LISA [6, 7] is a compiler generator based on attribute grammars, developed at University of Maribor at Slovenia. It generates a compiler and several other graphical tools from a textual AG specification, as can be seen in [8]. Since it generates a set of useful visual/graphical tools, it would be desirable to have a graphical way of specifying the AG too.

So the main idea is to enhance the front-end of LISA by developing a VPE, named VisualLISA, that assures the possibility to specify AGs visually, and to translate them into LISA specifications or, alternatively, into a universal XML representation designed to support generic AG specifications. The main objective of this environment is to diminish the difficulties regarding the specification of AGs not only in LISA but also for other similar systems.

The visual programming environment was automatically generated by DEViL, our choice among many other tools studied; so in this paper, the system is introduced and its use explained. However, our objective in this paper is not concerned with the discussion of compiler development tools, but show the benefits of using an effective one.

In Section 2, VisualLISA language and editor are informally described. In Section 3 the language is formally specified, defining syntactic rules, semantic constraints and translation scheme, which requires the description of the target language — in our case we present LISA’s syntax (the first target), and XAGra, an XML notation for AGs that we designed specially to support a universal representation.” In Section 4, the DEViL generator framework, used for the automatic generation of the visual editor, will be presented. In Section 5, following the informal conception and its formalization, using DEViL, the visual language and the editor implementation is shown. An overview
on how to use the editor to describe an AG, is given in Section 6 before concluding the paper in Section 7.

2 VisualLISA - A Domain Specific Visual Language

For many years we have been thinking and working with AGs. Inevitably we created an abstract mental representation of how it can be regarded and then sketched, for an easier comprehension and use. So we decided to implement a framework that follows that representation. The conception of that framework is described in this section.

2.1 The Language Conception

VisualLISA, as a new Domain Specific Visual Language (DSVL) for attribute grammar specification, shall have an attractive and comprehensible layout, besides the easiness of specifying the grammar model.

We think that a desirable way to draw an AG is to make it production oriented, and from there design each production as a tree. The Right-Hand Side (RHS) symbols should be connected, by means of a visible line, to the Left-Hand Side (LHS) symbol. The attributes should be connected to the respective symbols, using a connection line different from the one referred before, as both have different purposes (see Figure 1). The rules to compute the values of each attribute should exhibit the shape of a function with its arguments (input attributes) and its results (the output attributes). Two kinds of functions should be represented: the identity function (when we just want to copy values) or a generic function (for other kind of computations). Often a production has a considerable number of attributes and nontrivial computations. Therefore we think that for visualization purposes, the layout of each production should work as a reusable template to draw several computation rules. Hence, the rules are drawn separated from each other, but associated to a production.

Figure 1 depicts a simple example of the conception we made for the language. This example is based on a simple AG, called Students Grammar, used to process a list of students, described by their names and ages. The objective of the attribute grammar is to sum the ages of all the students. This grammar can be textually defined as shown in Listing 1.1. In Figure 1 just production P1 is shown. The attributes are associated to the symbols of the production. Moreover, the production has a semantic rule that computes the value of the LHS’s attribute, sum, by adding the value of the attributes in the RHS symbols, sum and age.

<table>
<thead>
<tr>
<th>Listing 1.1. Students Grammar</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1: Students → Student Students</td>
</tr>
<tr>
<td>{Students0.sum = Student.age + Students1.sum}</td>
</tr>
<tr>
<td>P2: Students → Student</td>
</tr>
<tr>
<td>{Students.sum = Student.age}</td>
</tr>
<tr>
<td>P3: Student → name age</td>
</tr>
<tr>
<td>{Student.age = age.value}</td>
</tr>
</tbody>
</table>
2.2 The Editor’s Main Features

VisualLISA editor should be compliant with the idea of offering a nice and non error-prone way of sketching the AG, as a first step; and an easy translation of the model into a target language, as a second step. So, three main features are highlighted: (i) syntax validation, (ii) semantics verification and (iii) code generation. The syntax validation restricts some spatial combinations among the icons of the language. In order to avoid syntactic mistakes, the edition should be syntax-directed. The semantics verification copes with the static and dynamic semantics of the language. Finally, the code generation feature generates code from the drawings sketched up. The target code would be LISAsl or XAGra. LISAsl specification generated is intended to be passed to LISA system in a straightforward step. XAGra specification generated is intended to give the system more versatility and further usage perspectives.

3 Specification of VisualLISA

The specification of VisualLISA bases on three main issues: i) the definition of the underlying language’s syntax; ii) the language semantics and iii) the description of the textual specifications into which the iconic compositions will be translated.

3.1 Syntax

The Picture Layout Grammar (PLG) formalism [9], is an attribute grammar to formally specify visual languages. It assumes the existence of pre-defined terminal symbols and a set of spatial relation operators. Our acquaintance with PLG formalism, from previous works, led us to use it to specify the syntax of VisualLISA. Listings 1.2 present some parts of the language specification.

Figure 2 shows the concrete and connector icons used for VisualLISA specifications. LeftSymbol is the LHS of a production, while NonTerminal and Terminal are used to compose the RHS. The second line of icons in Figure 2 presents the several classes of attributes. Function and Identity, both representing operations, are used to compute
the attribute values. The other icons connect the concrete symbols with each other, to rig up the AG.

### Listing 1.2. VisualLISA Partial Syntax Definition.

```
AG  →  contains(View, Root)
VIEW  →  labels(text, rectangle)
ROOT  →  left_to(PRODS, SPECS)
SPECS  →  contains(View
       over(LEXEMES, USER_FUNCS))
PRODS  →  group_of(SEMPRODS)
SEMPRODS  →  contains(View, left_to(
       group_of(group_of(RULE_ELEM))
       group_of(AG_ELEM))
AG_ELEM  →  LEFT_SYMBOL
        |  NON_TERMINAL
        |  TERMINAL
        |  SYNT_ATTRIBUTE
        |  TREE_BRANCH
        |  INT_ATTRIBUTE
        |  SYNT_CONNECTION
        |  INH_CONNECTION
        |  INT_CONNECTION

RULE_ELEM  →  FUNCTION
        |  IDENTITY
        |  FUNCTION_ARG
        |  FUNCTION_OUT
TERMINAL  →  label(text, rectangle)
INT_ATTRIBUTE  →  label(text, triangle)
INT_CONNECTION  →  points_from(
        point_to(dashline, ~INT_ATTRIBUTE), ~TERMINAL)
        |  point_to(arrow, ~FUNCTION)
        |  point_to(arrow, ~SYNT_ATTRIBUTE), ~FUNCTION)
```

Fig. 2. The Icons of VisualLISA

### 3.2 Semantics

In order to correctly specify an AG, many semantic constraints must hold. These constraints are related with the attribute values that depend on the context in which the associated symbols occur in a sentence. We separated these constraints into two major groups. One concerning the syntactic rules, Production Constraints (PC), and another the respective computation rules, Computation Rules Constraints (CRC).
The constraints concerning the VisualLISA’s semantic correctness, are not listed in this document, for the sake of the space. The complete set of constraints can be seen in [10].

3.3 Translation

The translation \( L_s \rightarrow \tau \rightarrow L_t \) is the transformation of a source language into a target language. \( \tau \) is a mapping between the productions of the \( L_s \) (VisualLISA) and the fragments of \( L_t \) (LISAsl \( \cup \) XAGra). These fragments will be specified in this sub-section.

A Context Free Grammar (CFG) is a formal and robust way of representing LISA specifications’ structure. Listing 1.3 presents that high-level CFG.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>p₁</td>
<td>LisaML ( \rightarrow ) language id { Body }</td>
</tr>
<tr>
<td>p₂</td>
<td>Body ( \rightarrow ) Lexicon Attributes Productions Methods</td>
</tr>
<tr>
<td>p₃</td>
<td>Lexicon ( \rightarrow ) lexicon { LexBody }</td>
</tr>
<tr>
<td>p₄</td>
<td>LexBody ( \rightarrow ) ( regName regExp ) *</td>
</tr>
<tr>
<td>p₅</td>
<td>Attributes ( \rightarrow ) attributes { type symbol . attName ; } *</td>
</tr>
<tr>
<td>p₆</td>
<td>Productions ( \rightarrow ) rule id { Declaration }</td>
</tr>
<tr>
<td>p₇</td>
<td>Derivation ( \rightarrow ) symbol ( \rightarrow ) Synbs ( \rightarrow ) compute { SemOperations }</td>
</tr>
<tr>
<td>p₈</td>
<td>Synbs ( \rightarrow ) symbols</td>
</tr>
<tr>
<td>p₉</td>
<td>epsilon</td>
</tr>
<tr>
<td>p₁₀</td>
<td>SemOperations ( \rightarrow ) symbol . attName = Operation</td>
</tr>
<tr>
<td>p₁₁</td>
<td>Operation ( \rightarrow )</td>
</tr>
<tr>
<td>p₁₂</td>
<td>Methods ( \rightarrow ) method id { javaDeclarations }</td>
</tr>
</tbody>
</table>

Reserved words, written in bold, indicate, in its majority, the beginning of important fragments. The fact of separating the structure in smaller chunks, makes the process of generating code easier and modular.

Regarding the literature, there is not an XML standard notation for AGs. So that, XAGra was defined using a schema. The whole structure of this schema can be separated into five big fragments: i) symbols — where the terminal, nonterminal and the start symbols are defined; ii) attributesDecl — where information about the attributes and the symbols to which they are associated is stored; iii) semanticProds — where the productions and the semantic rules are declared: in each production, the LHS, the RHS and the attribute computations are defined in a very modular way; iv) importations — where the modules or packages necessary to perform the computations are declared and v) functions — is the element where the user declare necessary functions. A more detailed explanation about these elements, their sub-elements and attributes can be seen in [10].

4 DEViL - A Tool for Automatic Generation of Visual Programming Environments

We searched for VPE generators like MetaEdit+ [11], but their commercial nature was not viable for an academic research. Also, we experimented VLDesk [12], Tiger [13], Atom³ [14] and other similar tools, however none of them gave us the flexibility that DEViL offered, as described below.

DEViL system generates editors for visual languages from high-level specifications. DEViL (Development Environment for Visual Languages) has been developed at
The University of Paderborn in Germany and is used in many nameable industrial and educational projects.

The editors generated by DEViL offer syntax-directed editing and all features of commonly used editors like multi-document environment, copy-and-paste, printing, save and load of examples. Usability of the generated editors and DEViL and can be found in [15]. DEViL is based on the compiler generator framework Eli [16], hence all of Eli’s features can be used as well. Specially the semantic analysis module can be used to verify a visual language instance and to produce a source-to-source translation.

To specify an editor in DEViL we have to define the semantic model of the visual language at first. It is defined by the domain specific language DEViL Structure Specification Language (DSSL) which is inspired by object-oriented languages and offers classes, inheritance, aggregation and the definition of attributes. The next specification step is to define a concrete graphical representation for the visual language. It is done by attaching so called visual patterns to the semantic model of the VL specified in DSSL. Classes and attributes of DSSL inherit from these visual patterns. Visual patterns [17] describe in what way parts of the syntax tree of the VL are represented graphically, e.g. we can model that some part should be represented as a “set” or as a “matrix”. DEViL offers a huge library of precoined patterns like formulae, lists, tables or image primitives. All visual patterns can be adapted through control attributes. E.g. we can define paddings or colors of all graphical primitives. Technically visual patterns are decorated to the syntax tree by specifying some easy inheritance rules in a DSL called LIDO.

To analyse the visual language, DEViL offers several ways. The first one results from the fact that editors generated by DEViL are syntax directed. Hence, the user cannot construct wrong instances of the VL it is limited by its syntax and cardinalities expressed in DSSL. Another way is to define check rules e.g. to check the range of an integer attribute or to do a simple name analysis on a name attribute. To navigate through the structure tree of the VL, DEViL offers so called path expressions which are inspired by XPath. They can be used in a small simple DSL to reach every node in the tree. After analysis, DEViL can generate code from the VL instance. This is done with the help of Eli which offers unparsers, template mechanism (Pattern-based Text Generator — PTG) and the well-known attribute evaluators from compiler construction.

5 Implementation of VisualLISA

After having all the requirements formally specified (Section 3) and a VPE generator chosen, the implementation of VisualLISA is a straightforward work and can be systematized in four main steps: i) Abstract Syntax Specification; ii) Interaction and Layout Definition; iii) Semantics Implementation; and iv) Code Generation.

5.1 Abstract Syntax

The specification of the abstract syntax of VisualLISA in DEViL, follows an object-oriented notation, as referred previously. This means that the nonterminal symbols of the grammar are defined modularly: the symbols can be seen as classes and the attributes of the symbols as class attributes.
The syntax of the visual language is determined by the relations among their symbols. Therefore, for an high level representation of the language’s syntax, a class diagram can be used. This diagram should meet the structure of the PLG model in Figure 1.2. The final specification for the language is then an easy manual process of converting the diagram into DSSL. Figure 3 shows a small example of the diagram and the resultant specification.

There are two types of classes in this notation: concrete and abstract. The concrete classes are used to produce a syntax tree, which is manipulated in the other steps of the environment implementation. The abstract classes are used to group concrete classes with the purpose of defining syntactic constraints. These classes generate the syntax-directed editor.

In order to make possible the specification of separated computation rules reusing the same layout of a production, we used DEViL’s concept of coupled structures [18]. It couples the syntactic structure of two structure tree — for VisualLISA we used the structure of symbol Semprod, which is used to model a production. In practice, it means that the layout defined for a production is replicated whenever a computation rule is defined, maintaining both models synchronized all the time.

5.2 Interaction and Layout

The implementation of this part, in DEViL, consists in the definition of views. A view can be seen as a window with a dock and an editing area where the language icons are used to specify the drawing.

VisualLISA Editor is based in four views: rootView, to create a list of productions; prodsView, to model the production layout; rulesView, to specify the semantic rules reusing the production layout and defsView, to declare global definitions of the grammar.

At first the buttons of the dock, used to drag structure-objects into the edition area, are defined. Then the visual shape of the symbols of the grammar for the respective view are defined. Figure 4 shows parts of view definitions and the respective results in the editor. The code on the left side of Figure 4 defines the view, the buttons and the behavior of the buttons. The default action is the insertion of a symbol in the editing
area, but it can be extended. The bluish rectangular image represents the button resultant from that code.

```plaintext
1 VIEW rootView ROOT Root{
2 BUTTON IMAGE "img : btnProd"
3 INSERTS Semprod
4 INFO "Production";
5 }
```

**Fig. 4.** Parts of View Definitions and Respective Visual Outcomes

Symbol NonTerminal is represented by the orange oval in Figure 4. The code on the right reveals the semantic computation to define the shape of that symbol. Shape and other visual aspects of the tree-grammar symbols are automatically defined associating, by inheritance, visual patterns.

### 5.3 Semantics
As long as VisualLISA is defined by an AG, the contextual conditions could be checked using the traditional approach. DEViL is very flexible and offers some other ways to implement this verification module. The approach used to develop VisualLISA, is completely focused on the contexts of the generated syntax tree. DEViL offers a tree-walker, that traverses the tree and for a given context — a symbol of that tree — executes a verification code, returning an error whenever it occurs. With this approach it is easy to define data-structures helping the verification process. This approach is very similar to the generic AG approach, but instead of attributes and semantic rules, it uses variables which are assigned by the result of queries on the tree of the model.

Listing 1.4 shows the code for the implementation of a constraint defined in [10].

**Listing 1.4.** Implementation of Constraint: “Every NonTerminal specified in the grammar must be root of one production”

```plaintext
checkout addCheck Semprod {
1 set n [length {c : getList {obj.grammarElements CHILDREN[LftSymbol]}}]
2 set symbName [c : get {obj.name.VALUE}]
3 if {n == 0} {
4 return "Production 'symbName' must have one Root symbol!"
5 } elseif {n > 1} {
6 return "Production 'symbName' must have only one Root symbol!"
7 } return ""
8 }
```

A considerable amount of the constraints defined in Section 3.2 were verified resorting to the Identifier Table, which is a well known strategy in language processing for that purpose.

### 5.4 Code Generation
The last step of the implementation, concerning the translation of the visual AG into LISA or XAGra, can be done using the AG underlying the visual language (as usual
in language processing). For this task, DEViL supports i) powerful mechanisms to ease the semantic rules definition; ii) facilities for extend the semantic rules by using functions and iii) template language (PTG of Eli system) incorporation to structure out the output code.

The use of patterns (templates) is not mandatory. But, as seen in the formal definition of LISA and XAGra notation (Section 3.3), both of them have static parts which do not vary from specification to specification. Hence templates are very handy here. Even with templates, the translation of the visual AG into text is not an easy task. Some problems arise from the fact that there is not a notion of order in a visual specification. We used auxiliary functions to sort the RHS symbols by regarding their disposition over an imaginary X-axe. Based on this approach we also solved issues like the numbering of repeated symbols in the production definition.

The templates (invoked like functions) and the auxiliary functions, together with LIDO-specific entities like INCLUDING or CONSTITUENTS, were assembled into semantic rules. These entities collect the values of attributes from different productions at the same time. Then these values are used by templates or functions in order to define the translation module. One module was defined for each target notation. New translation modules can be added, for supporting new target notations.

6 AG Specification in VisualLISA

Figure 5 shows the editor appearance, presenting the four views of our editor.
To specify an attribute grammar the user starts by declaring the productions (in *rootView*) and rigging them up by dragging the symbols from the dock to the editing area (in *prodsView*), as commonly done in VPEs. The combination of the symbols is almost automatic, since the editing is syntax-directed. When the production is specified, and the attributes are already attached to the symbols, the next step is to define the computation rules. Once again, the user drags the symbols from the dock, in *rulesView*, to the editing area, and compounds the computations by linking attributes to each other using functions. Sometimes it is necessary to resort to user-defined functions that should be described in *defsView*. In addition, he can import packages, define new data-types or define global lexemes.

The AG example addressed in Section 2.1 is composed of three productions with associated semantic rules. Production P1 is already shown in Figure 1. The other two productions are shown in Figure 6.

![Fig. 6. Specification of Productions P2 and P3, (a) and (b) respectively, with associated semantic rules.](image)

In production P2, the identity function is used to copy the value of the attribute *age* to the attribute *sum*. Production P3 makes use of terminal symbols and associated intrinsic values. The computation rule, in this production, is based on the conversion of the textual value of the age into an integer.

When the grammar is completely specified and semantically correct, code can be generated. Listing 1.5 shows, in LISA and *XAGra* notations, the code generated for production P1 in Figure 1.

**Listing 1.5. Code Generated for a) LISA and b) XAGra**

```plaintext
1 2 a)
3   rule Students {} {
4       STUDENTS := STUDENT STUDENTS compute {
5         STUDENTS . sum = STUDENTS[1] . sum + STUDENT . age;
6     };
7   }
```

7 Conclusion

After many years working in specification and implementation of compilers supported by Attribute Grammars, it became clear that a modular and reusable approach to AG development is highly recommendable and necessary. On the other hand, the work on program comprehension tools emphasized the importance of software/data visualization. The combination of those two areas of R&D with a third one, the development of Visual Languages, gave rise to the proposal of creating a VL for AGs. The obligation to write text-based AG specifications imposed by several compiler generator tools and the habitual way of sketching AGs on paper in the form of a decorated tree, shortening the gap to the mental representation of an AG, reinforced the appropriateness of that proposal.

In this paper we introduced VisualLISA, a new Domain Specific Visual Language (DSVL), which enables the specification of AGs in a visual manner and the translation of that visual AG into LISA or XAGra (an XML notation to support generic AG specifications). We were mainly concerned with the design of the language and its formal and automatic implementation. In this phase of our project we neither focused on the usability of the language nor on its scalability. We focused on the specification, aiming at showing the formal work behind the visual outcome, and on the implementation of the underlying environment to specify AGs. At this point we highlighted the use of DEViL in order to create the desired environment, through a systematic approach of development. Also, an example was presented to show the steps to build an AG with VisualLISA.

In the future, it is our objective to perform at least, two experimental studies involving VisualLISA: one to assess the usability of the language regarding the visual vs textual approaches for developing AGs; and another one to test the scalability of the language and environment, regarding the hypothesis that it was created to cope with small AGs. We are also interested in assessing the comprehension of AGs; maybe VisualLISA would be very handy on this matter, working as AGs visualizer.
References