Abstract. To be a debugger is a good thing! Since the very beginning of the programming activity, debuggers are the most important and widely used tools after editors and compilers; we completely recognize their importance for software development and testing. Debuggers work at machine level, after the compilation of the source program; they deal with assembly, or binary-code, and are mainly data structure inspectors. Alma is a program animator based on its abstract representation. The main idea is to show the algorithm being implemented by the program, independently from the language used to implement it. To say that ALMA is a debugger, with no value added, is not true! ALMA is a source code inspector but it deals with programming concepts instead of machine code. This makes possible to understand the source program at a conceptual level, and not only to fix run time errors. In this paper we compare our visualizer/animator system, ALMA, with one of the most well-known and used debuggers, the graphical version of GDB, the DDD program. The aim of the paper is twofold: the immediate objective is to prove that ALMA provides new features that are not usually offered by debuggers; the main contribution is to recall the concepts of debugger and animator, and clarify the role of both tools in the field of program understanding, or program comprehension.

Keywords: Program Comprehension, Program Animation, Debugger, Alma, DDD.

1. Introduction

In the context of our research on language-based tools and program understanding, we developed, some time ago, a program animator, ALMA, which has been confronted many times against debuggers; a common question is: “what is the value added by ALMA to the usefulness of traditional debuggers?”

ALMA was conceived in order to prove the potential of using a traditional compilers' approach to enable the construction of generic (algorithm/language independent) tools for program visualization/animation. That developing effort
was justified because we did not find in literature any tool of that kind; after reviewing the area of program animation, we have concluded that a generic visualize animator (not specific for a family of algorithms or for a given programming language) was lacking. So it is difficult to compare ALMA with similar tools.

After attaining that methodological/technological objective, it became evident that ALMA could also be used to help programmers and software-engineers in program understanding. This is possible because ALMA let us abstract from the concrete syntax of programming languages and aids in associating (operational) meaning to programming concepts.

Our intention in this article is to claim and justify that ALMA is not a debugger! We state the differences between a debugger (code analysis) and a generic program animator (algorithm analysis). We want to prove that the animators offer different features to the user. So, ALMA can be a good help in pedagogical scenarios as well as for program comprehension.

We start the paper, on one hand writing about debuggers (section 2) and introducing (subsection 2.1) one in particular (DDD), and on the other hand presenting ALMA (subsection 3.1) in the context of animators (section 3).

Then, in section 4, we compare ALMA versus DDD, presenting general arguments, and going through a concrete example --- the analysis (i.e., the visual inspection) of the factorial program to comprehend the recursive function-call mechanism. To compare both tools, ALMA and DDD, we must study and confront many details. However, in this paper, we will focus just on those topics that we consider the most important: the purpose of the tool; its architecture; how and what is visualized.

### 2. Debuggers

In this section we intend to characterize the class of programs designated by debuggers. As we can find in literature ([17, 14]) lots of things had been already written about debuggers. Instead of trying our own definition and characterization, we will present well-known and accepted descriptions.

According to many authors, and also found in Webopedia\(^1\), a first and very simple definition of debugger says that it is a program used to find errors (bugs) in other programs, allowing the user to stop the execution of a given program at any point and examine and change the values of variables. So, it is possible to list the three most important characteristics of a debugger:

---

ALMA versus DDD

- Helps a programmer to find bugs in programs;
- Allows to inspect variable values;
- Allows the use of breakpoints and step-by-step execution.

Wikipedia\(^2\) presents a more sophisticated and complete definition that corroborates the previous one:

“A debugger is a computer program that is used to test and debug other programs. . . 
When the program crashes, the debugger shows the position in the original code if it is a source-level debugger or symbolic debugger, commonly seen in integrated development environments. If it is a low-level debugger or a machine-language debugger it shows the line in the disassembly. . . 
Typically, debuggers also offer more sophisticated functions such as running a program step by step (single-stepping), stopping (breaking) (pausing the program to examine the current state) at some kind of event by means of breakpoint, and tracking the values of some variables. . . 
Some debuggers have the ability to modify the state of the program while it is running, rather than merely to observe it”.

The last statement emphasizes the important functionality that should be offered by any debugger concerning the possibility to handle the set of variable-value pairs. Many programmers (especially those accustomed to IDE for software development) don't like to work with console debuggers. Nowadays, they use more sophisticated tools with navigation, visualization and animation features. As we can read in [14]:

“Visualization of computer algorithms, data structures, and program execution is a specific area in which visual tools can be used effectively to enhance the productivity of software development and usage. A visual debugger gives the user an opportunity to interact with the program graphically. The graphical display enhances the understanding of the specific way in which the program proceeds.”

Since ALMA is an animation system, it should be compared with modern debuggers with a visual and windows-based interface.

GVD\textsuperscript{3}, the GNU Visual Debugger now a component of GNAT (the GNU Ada IDE), could be an interesting choice. However, we chose DDD (Data Display Debugger) because it is one of the more sophisticated and well-known (more used) members of that kind of recent debuggers; an introduction to that tool, that will be further compared with ALMA (section 4), is presented in the next subsection.

2.1. DDD at a glance

The standard Linux debugger is GDB \cite{9}, the GNU debugger. GDB provides an interactive text-base method for accomplishing the tasks referred above, including step-by-step execution, breakpoints, variable watch and other options that are expected in a full-edged debugger. GDB is in fact a tool still under development.

The last novelty relies on adding "reversible debugging" support \cite{4} --- allowing a debugging session to step backwards, much like rewinding a crashed program to see what happened.

To make easier the debugging of a source program, DDD \cite{17} rise up as an X-Windows front-end to both GDB and DBX \cite{7} debuggers. DDD is just a front-end --- it does not do any debugging. Instead, it sends all user commands to a GDB (or DBX) running process, and limits itself to display the answer received from that back-end process.

Besides the usual FE features, such as the handling set and display of breakpoints, DDD provides an interactive graphical data display, where data structures are shown as graphs. So, using DDD, we can follow the program execution visualizing the source code line-by-line and also watch its data structures.

Hence the DDD depends on GDB, and GDB depends on GCC, the Gnu Compiler Collection (previously, meaning Gnu C Compiler)\textsuperscript{4}, it is relevant to understand GCC's architecture, which is illustrated in Fig. 1.


\textsuperscript{4} Remember that to run GDB on a program, it is necessary to compile before that program with GCC with debugging option activated.
Looking to the architecture of GCC we can remark that: the FE reads the source code and builds a parse tree; the parse tree is used to generate an RTL\(^5\) instruction list based on the named instruction patterns; and the instruction list is matched against RTL templates to produce assembly or machine-code.

This architecture allow us to understand how GCC supports multiple languages --- like Ada, Bash, C, C++, Chill, Fortran, Java, Modula, Pascal, Perl and Python; in fact, the FE parses each one and converts to RTL. Since RTL is a common intermediate representation (IR), very close to assembly language, this language is used by GCC to provide support to different languages. We additionally can say that to port the compiler to a new language is just needed to modify the FE; to alter the target architecture implies just a change on the BE.

\(^5\) Register Transfer Language.
Based on that kind of hierarchy (GCC $\rightarrow$ GDB $\rightarrow$ DDD), DDD is able to debug programs written in the languages supported by GCC (listed above). It supports: machine-level debugging; hypertext source navigation and lookup; breakpoint, back trace, and history editors; program execution in terminal emulator window; debugging on remote host; GDB/DBX/XDB command-line interface with full editing, history, and completion capabilities.

As can be seen in Fig. 3, DDD environment consists of four windows:
- The “debugger window” which contains the actual communication between DDD and GDB;
- The “source window” which contains the source program and the basic source debugging action;
- The “command tool window” which contains buttons to activate most of the debugging actions (stepping up and down, setting breakpoints, etc);
- The “data window” which contains all data-related information, such as variable and function watching.

With all these powerful features, the debugging of a program becomes an easy task.

3. Animators

The role of the visualization technology, in fields like program comprehension and software engineering, is strongly recognized by the computer science community as a very fruitful feature. Indeed, the use of software visualization functionalities allows us to get a high quantity of information in a faster way.

The purpose of an animation system is to construct (automatically) program visualizations in order to help the programmer to inspect data and control flow of a source program and to understand its behavior. This kind of information can be used by the programmer to recognize: the main tasks performed in the program; the changes in the value of variables; or possible errors. It also can be used to verify the correctness of the program results and their meaning.

The values computed by a program have an effect over the object under control; so it will be also desirable to visualize the program effects, this is, the meaning of the program.

In an animation system, a special care should be taken with content (subset of program information that will be visualized). It is also necessary to decide: how this content will be visualized; what kind of visualizations will be used; what kind of interaction facilities should be made available; how to solve scalability problems; the colors/sound usage, etc.

There are also several kinds of views that can be produced: operational (program level) or behavioral (problem level); static (compile-time) or dynamic (runtime); structural or quantitative (based on metrics or other kind of statistical information).
Animation systems usually are designed to include some features that are not available in debuggers. While animation systems are strongly oriented towards program comprehension --- and so they care about high-level/abstract concepts related with cognitive models --- traditional debuggers are concerned with low-level/concrete concepts focused on the machine comprehension.

Usually, in animation systems, the main research concerns are visualization techniques: to adapt to the knowledge level of the user; to provide multiple views; to map one view to another; to support learner-built visualizations; to complement visualizations with explanations; to support different abstraction levels, etc. These features are essentially concept-oriented. In debuggers, the concerns are: how to enable forward or backward traversal of the execution path; how to report historical information; how to support user’s input data; how to report one variable life cycle; how to use dependency graphs, etc. These kinds of features are essentially operation-oriented.

In the next subsection we present the ALMA system, explaining briefly its architecture, how it supports different source languages, and the animation process.

3.1. Alma at a glance

After reviewing the existing systems | like BALSA([2] in [12]), TANGO ([13], PECAN ([10] in [12]), FIELD ([11] in [12]), JELIOT [5] | we conclude that almost all were constructed for specific languages (language dependent) and use techniques like code instrumentation or special data types that imply changes in the source code. We also found systems, like CENTAUR [1], which can be used to construct program animators but are based on semantic annotations. We decided to design ALMA to be a generic tool for program visualization and animation based on the internal representation of the input program in order to avoid any kind of source code annotation, and to be able to cope with different programming languages.

To fulfill such requirements, we were inspired by the classic structure of a compiler and we conceived an architecture that separates the source program recognition from its animation, using a decorated abstract syntax tree (DAST) as internal representation (see Fig. 2).
ALMA was implemented in Java, using and reusing the compiler generator system LISA [8], as specified and described in [15], [6] and [16].

ALMA system has a front-end (FE) specific for each language and a generic backend (BE). It uses a decorated abstract syntax tree (DAST) for the internal representation of the program's meaning; this DAST is the connection between the FE and the BE. Using a DAST as an intermediate representation, we isolate all the source language dependencies in the FE, while keeping the generic animation engine in the BE. The DAST is built using a set of pattern rules [3]. A pattern represents an abstraction of a programming concept. Each one of these patterns is composed by two parts: a structural component (given by a grammar production) and a semantic component (given by a set of attribute occurrences affected to the symbol that labels each tree-node).

These patterns capture the abstract syntax of each entity (value or operation) in order to preserve and keep, via attributes, the necessary information to express static semantics of those values or operations. So the DAST do not reflect directly the source language syntax, and in this case, the tree generated by the FE (after parsing the source) represents the input program.

Applying specific rewrite rules (which are used according to the pattern-tree found in the DAST) to the execution tree, we obtain a description of the different program states, simulating its execution.

A Tree-Walker Visualizer, traversing the execution tree and applying visual rules, creates a representation for the nodes, generating a visualization of the program tree in that moment. Then the DAST is rewritten (to obtain the next internal state), and redrawn, generating a new visualization which reflects the new state of the program.

This approach, using a DAST as internal representation and a set of pattern rules allows us to easily construct different abstraction levels of the same program since the operational view till the behavioral view. For that, it suffices to associate a new set of rewrite and visualize rules to the DAST patterns. This system is based on the concepts involved in a program and not directly in the source code.
As can be seen in Fig. 5, ALMA's interface is split in four windows --- 3 main windows and 1 with the buttons for navigation --- with the following content:
- The identifier table (on the top, to the left);
- The source text of the program to be animated (on the bottom, to the left);
- The program tree (to the right, occupying most of the display);

The interaction buttons that allow the user to control the system. Using back and forward controls the user will be able to navigate through the program animation, step by step.

Colors are used to make easier to track the animation of a program: Red color indicates the identifier / subtree that will be changed by the next execution step (the execution of the program-statement also colored in red); Green color indicates the new status of the program (subtree /identifier) after the last change.

In the next section we finally show how different is ALMA from DDD. For that purpose a simple source program written in C --- the recursive implementation of the factorial function --- is used. This example, although simple, is short and allow us to focus in the most important and crucial details.

### 4. Alma vs DDD

To illustrate how we perform this study, we discuss, in this section, two specific cases: recursive function invocation (section 4.1) and data types representation (section 4.2).

Before that, we synthesize in table 1 the criteria used for the global comparison. Table 1 is also fulfilled with final results of the referred comparison.

**Table 1. Criteria for comparing Alma vs DDD**

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Alma</th>
<th>DDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>GUI</td>
<td>windows based</td>
<td>windows based</td>
</tr>
<tr>
<td>Breakpoints</td>
<td>not allowed</td>
<td>allowed</td>
</tr>
<tr>
<td>Stepping</td>
<td>back and forward</td>
<td>back and forward</td>
</tr>
<tr>
<td>Identifiers table</td>
<td>shown</td>
<td>shown</td>
</tr>
<tr>
<td>Registers and memory addr</td>
<td>not shown</td>
<td>not shown</td>
</tr>
<tr>
<td>Code</td>
<td>simulated</td>
<td>simulated</td>
</tr>
<tr>
<td>Structured variables</td>
<td>values shown</td>
<td>values shown</td>
</tr>
<tr>
<td>Function call</td>
<td>separated windows</td>
<td>separated windows</td>
</tr>
<tr>
<td>Parameter passing</td>
<td>shown</td>
<td>shown</td>
</tr>
<tr>
<td>Source language</td>
<td>independent</td>
<td>independent</td>
</tr>
<tr>
<td>Machine language</td>
<td>source</td>
<td>source</td>
</tr>
<tr>
<td>Abstraction level</td>
<td>source-code tree</td>
<td>memory mapping</td>
</tr>
<tr>
<td>Program Visualization</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Program Animation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.1. Recursive function invocation

To do the comparison between ALMA and DDD, we will show how each one of the tools holds and shows the information for the same source program listed below --- a C program that implements the factorial function.

After loading *factorial program*, DDD displays the information illustrated in Fig. 3.

```c
#include <stdio.h>
int factorial(int n) {
  int res=1;
  if (n>0) {
    res = n*factorial(n-1);
  }
  return res;
}

int main() {
  int a, r;
  scanf("%d",&a);
  r = factorial(a);
  printf("%d ",r);
  return 0;
}
```

Our main purpose, exploring this example, is to focus on the recursive mechanism: How does DDD deal with the function-call process? How does it cope with the parameter-passing process?

Suppose that we add a breakpoint to the program at line 5 (recursive call). Fig. 3 and 4 illustrate what happens when, evaluating the factorial of 4, we execute 2 steps (selecting Step action twice), after DDD stops at the breakpoint. Observing those figures, we see that: in the stack the double call of factorial function is explicit, first with the value 4, and then with the value 3; the next line executed is the line 4; and this is also visible in the frame and in the registers table.

But, if we are trying to understand which values are actually passed and how they are assigned to the function parameters, DDD is not of a big help, because it just displays the values of the arguments when the function execution starts (after the assignment).
If we do not know assembly language at all, the use of a debugger is very complex or not really helpful. In fact, if we are just interested in understanding the program's control or data flow, the visualization of that mess of registers, and hexadecimal codes or addresses can be awkward. If this is the case, the use of ALMA can be helpful. Following the previous procedure (above done
for DDD), we can observe in figures 5, 6, and 7 the output of ALMA for the function-call, parameter passing, and function execution.

Each time a function is called a new window is opened, being drawn in a cascade style (see Fig. 6). As the new function has its own memory area, a new window shows the function execution environment (IdTable), i.e., the local and global identifiers; the program-tree for that function is also displayed.

Looking at Fig. 6, we see that the call of \textit{factorial} function causes a jump to the function-header to show how the parameters are passed --- this is a completely different approach of the one followed by DDD. Notice that the identifier table, during the argument-passing phase, is different from the identifier table available during the execution of the function body (Fig. 7). At this moment (argument passing), the active identifier table is the one belonging to the calling function plus the function arguments (defined in the \textit{factorial} header).

After passing the parameters to the function, a new sub-window --- named \texttt{Executing function} --- opens at the right of the previous one to display the tree corresponding to the function body (Fig. 7) and to show its execution. At this phase, the identifier table is the one strictly accessible inside the called function (including just the global variables and those variables local to \textit{factorial}).

It is evident that ALMA displays more information, and in a clearer way, than DDD for the same execution process: the function/procedure call (recursive or not).

Fig. 5. Executing the main function in ALMA --- reading a variable
4.2. Data types representation

Another interesting/important case to compare DDD and ALMA is the visual inspection of structured data types. For that purpose, we will consider the visualization of arrays.
In DDD, displayed pointer values are dereferenced by a simple mouse click, allowing unfolding arbitrary data structures interactively. DDD automatically detects if multiple pointers refer to the same address and adjusts the display accordingly. DDD has a major drawback: each and every pointer of a data structure must be dereferenced manually. While this allows the programmer to set a focus on specific structures, it is tedious to access, let us say, the 100th element in a linked list.

In ALMA, each element of a structured data type is displayed in a row of a table, freeing the user from the tedious task of clicking in each element to discover the next one. If the array is larger, the user only needs to scroll down. These situations are illustrated in Fig. 8 and 9 for DDD and ALMA, respectively.

So, we are able to say, after comparing DDD and ALMA tools, that ALMA provides a more effective aid in the context of program comprehension/understanding.
Below are the stronger arguments that came out from the comparative study, to affirm that definitely ALMA is not a Debugger.

- ALMA is not a traditional (low-level) debugger in the common sense of a tool that helps the programmer to fix errors in the assembly/machine-code. Alma creates visualization for abstract programming concepts and animates them. ALMA can simulate the source code execution (on an operational visualization level) but it does not compile the source program, neither it reuses the assembly code.

- ALMA is not C/RTL-oriented, it is language-independent. Even knowing that DDD can cope with any language L that compiles to RTL it is much more difficult/complex to create a compiler to RTL then a L -- FE for ALMA (mapping L-statements to abstract high-level concepts).

- DDD works at an operational-level (machine level), while ALMA can produce visualizations at a concept-level (programming abstract level) or at an operational level as we presented in this paper. It depends on the visualization rules associated to the program internal representation.

- ALMA is aimed at explaining (illustating) the program semantics that can be useful to detect/understand algorithmic errors, while DDD and other debuggers are oriented to low-level error detection.

Notice the way ALMA deals with parameter passing referred above. In any high-level language this mechanism is entirely embedded in the subprogram invocation, and the underlying idea of assigning values (the actual parameters) to the local variables (the formal parameters) is not explicit anywhere. As long as ALMA works with concepts, this mechanism is made explicit...
Of course ALMA has some limitations when compared with most debuggers --- ALMA is an academic prototype, while the others are professional, or commercial, tools. ALMA does not support, at this moment, *pointers* neither *objects*, and is much poor concerning the interaction with the user during the animation (no breakpoints available, etc.). To deal with multiple language complex applications, although not experimented, we defend that it will not be a problem, as it is a matter of converting the various modules to the DAST uniform representation.

5. Conclusion

ALMA and DDD (or any other visual debugger) are very similar, at least at a first sight, and many people says that ALMA is just a new debugger without any value added. Because of that we decided to carry on a study aiming at gathering arguments to agree or refute that statement. For that purpose, we selected DDD to confront with ALMA. This article reports that work, concluding that ALMA adds value to a debugger! After a careful study of both tools (just summarized here), we went through an example --- the recursive computation of the factorial --- highlighting the points of the visual inspection offered by each tool in order to support the arguments presented at the end of section 4.

6. References

ALMA versus DDD


Daniela da Cruz received a degree in "Mathematics and Computer Science", at University of Minho, and now she is starting a Ph.D. degree in "Computer Science" also at University of Minho, under the MAPI doctoral program. She joined the research and teaching team of "gEPL, the Language Processing group" in 2005. She is teaching assistant in different courses in the area of Compilers and Formal Development of Language Processors; and Programming Languages and Paradigms (Procedural, Logic, and OO). As a researcher of gEPL, Daniela is working with the development of compilers based on attribute grammars and automatic generation tools. She developed a compiler and a virtual machine for the LISS language (an imperative and powerful programming language conceived at UM). She was also involved in the PCVIA (Program Comprehension by Visual Inspection and Animation), a FCT funded national research project; in that context, Daniela worked in the implementation of "Alma", a program visualizer and animator tool for program understanding. She is now enrolled in a new bilateral cooperation project with Slovenia under the subject "Program Comprehension for Domain Specific Languages".
Maria João Varanda Pereira received the M.Sc. and Ph.D. degrees in computer science from the University of Minho in 1996 and 2003 respectively. She is currently an Adjunct Professor at the Polytechnic Institute of Bragança in the Informatics and Communications Department. Her research interests include programming languages, compilers, grammar-based systems, visual languages, program comprehension, animation systems and domain specific languages. She was responsible for PCVIA (Program Comprehension by Visual Inspection and Animation), a FCT funded national research project. She is involved in bilateral cooperation projects with Slovenia since 2000.

Pedro Rangel Henriques got a degree in "Electrotechnical/Electronics Engineering", at FEUP (Oporto University), and finished a Ph.D. thesis in "Formal Languages and Attribute Grammars" at University of Minho. In 1981 he joined the Computer Science Department of University of Minho, where he is a teacher/researcher. Since 1995 he is the coordinator of the "Language Processing group". He teaches many different courses under the broader area of programming: Programming Languages and Paradigms (Procedural, Logic, Functional and OO); Compilers and Formal Development of Language Processors; etc. He is co-author of the "XML & XSL: da teoria à prática" book, publish by FCA in 2002. Pedro Rangel Henriques has supervised M.Sc. (16) and Ph.D. (14) thesis, and more than 100 graduating trainingships/projects, in the areas of: language processing (textual and visual), and structured document processing; program animation and program comprehension; knowledge discovery from databases, data-mining, and data-cleaning. He also was responsible for several applicational projects (in the interface university/external-community, industry), mainly in the area of information systems (databases and web oriented). From 2002 until 2004 he was the Head of the Department, and at moment he is the President of APPIA, the Portuguese Association for Artificial Intelligence.

Received: July 16, 2008; Accepted: December 3, 2008.