**A UML Virtual Machine based on AST execution and its environment**

Software Engineer Bruno Buzzi Brassesco – Reviewer VM expert Eliot Miranda

[[1]](#footnote-1)

# Abstract

In a big project for a government (in 2009) with 10 functional analyst (from company A) and 20 programmers (from company B) to implement a system with 400 classes or so. By contract the functional analysts (company A) have to produce an UML Model (in Enterprise Architect) and based on the UML Model create Use Cases (as Word Documents) to be sent to the programmers (the software factory - company B). Also a lot of requirements needed to be broke down to be programmed. So some modules were in programming phase while others were being created or broken down.

This methodology had a lot of problems, most Use Cases were rejected by the Software Factory because different errors in the UML model. The most common errors were: referencing a missing class, a missing attribute, incorrect attribute type, missing multiplicity, incompability with other part of the model (when more than one functional analyst were responsible for a Use Case) to name just a few. This situation was translated into higher cost and development time.

Translate a Word document into code was an expensive and time consuming task. The gap between end user, functional analyst and programmers was too wide.

In order to improve this situation it was visualized that another tool was needed in the development process.

In this paper is detailed the tool developed to solve different challenges derived from above context. The whole problem should be solved incrementally then most critical issues must be confronted first.

# Introduction

The first version of the tool was a simple application to import XMI files from Enterprise Architect and display classes, relations and attributes to detect missing UML elements. The rejection of Uses Cases felt dramatically (around ~70%).

The second version allowed the creation of UML instances in order to dynamically visualize how elements interact with each other. This helped to refine the UML Class diagram and improve the Uses Cases creation.

The third version added a web interface to show UML instances as a Web Application enabling a much better communication with end users to break down requirements.

The fourth version introduced a major change, the support for method execution including Blocks. With this version a Use Case could be executed before been written as Word Document. This assured 100% of accuracy in Uses Cases and saved a lot of time and money.

This last version can be truly called an UML Virtual Machine (that was named U-Fabrik).

U-Fabrik was implemented in Dolphin Smalltalk 6 in 2009 and it was migrated to Dolphin Smalltalk 7.0 in 2018.

An UML Virtual Machine (UVM) is a program that executes code based on an UML Class diagram (and an UML State diagram[[2]](#footnote-2)). The code to execute can have any representation like bytecodes, AST or other. The UVM analyze the code representation and creates an execution context in order to perform the code execution.

The UVM has to have the ability to create class instances based on the UML Class diagram and set and get instances’ attributes. Also it has to handle in some way the instantiation of UML Class relations (this UVM implement instances creation and update through UVM primitives and traits).

The code execution performed by the UVM has different types of restrictions defined on the UML Class diagram. For example it is forbidden to assign a string value to an integer attribute of a class.

Classes’ relations are defined by multiplicity on the UML Class diagram. Class A may have a 1:N relation with class B then an instance of class A it will have N instances of class B and each class B it will have only an instance of class A. The UVM has to do some type checks when class instances are being composed. These type checks has to include inheritance of the model too.

To summarize the characteristics of a UVM:

1. A way to import UML Class diagrams.
2. Ability to instantiate the UML Class diagram.
3. Basic building blocks to create complex operations in 4.
4. Ability to create code and code representation.
5. Ability to create context execution to perform code execution.

In order to support these five points it is very desirable to have a development environment to perform each activity.

# UML Virtual Machine

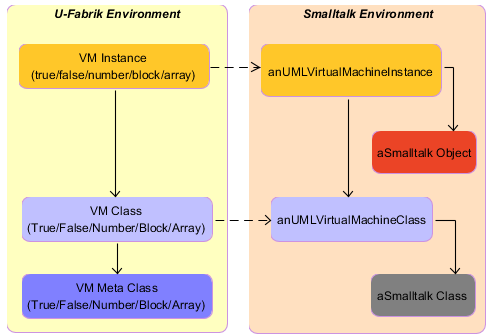
In the present case the implemented UVM comply with these five points and there is also a development environment to create simulations.

1. UML Class diagrams are imported through XMI 2.1 interface used by any UML Case tool such as Visual Paradigm, Enterprise Architect and so on ...
2. The UML Class diagram instantiation is supported by a powerful Meta Model.
3. The UVM has a basic hierarchy of classes to allow the creation of much more complex functionality.
4. The code representation is done with an AST (abstract syntax tree) using a Smalltalk Parser to transform a text into an AST.
5. Execution Context is implemented as Smalltalk classes and it’s at the core of the UVM.

The code can be found at:

<https://github.com/brunobuzzi/U-Fabrik>

# U-Fabrik Architecture

The following image depict a general architecture of the whole U-Fabrik environment including the UML Virtual Machine.

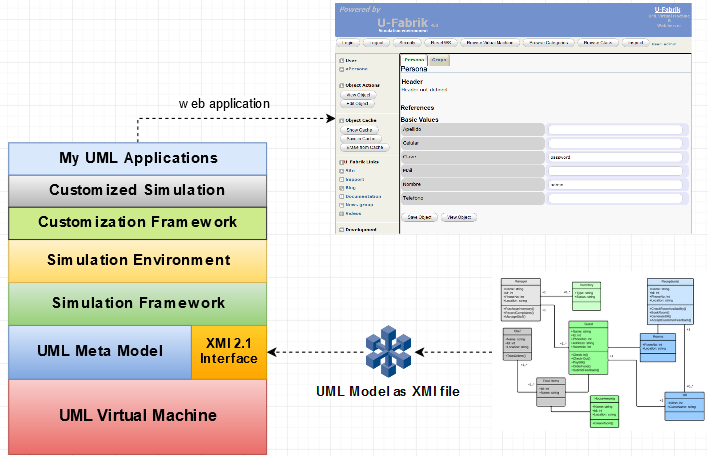


Image 1 - VM Objects representation in both environments

Image 2 - U-Fabrik General Architecture

The XMI 2.1 interface used to import XMI files from Enterprise Architect®, Visual Paradigm®, and so on. The Meta Model is used to represent all elements contained in the XMI file. The UML Virtual Machine simulate behavior through message sending. On the top of the UML Virtual Machine there is a Graphical Environment for prototyping and simulation. The default web prototype is done automatically (after the XMI installation). The Customization Framework combines the Simulation Framework and Simulation Environment to achieve customization through method implementations and drag & drop operations (see web simulation chapter).

When a XMI file is imported each element of the UML diagram is instantiated as part of the UML Meta Model. The UVM is capable to use the elements of the Meta Model to execute basic and complex behavior. The Graphical Environment allow to create new methods and from these new methods the web application can be fully customized (see chapter web simulation).

# U-Fabrik Virtual Machine (UVM)

In U-Fabrik Virtual Machine objects are wrapped around Smalltalk objects. Each UVM object is an instance of Smalltalk class UMLVirtualMachineInstance. This class has an instance variable named “smalltalkObject” which holds the actual Smalltalk object.

All basic classes of the U-Fabrik Virtual Machine are instances of UMLVirtualMachineClass and are instantiated as instances of UMLVirtualMachineInstance (Smalltalk class) where “smalltalkObject” points to the actual Smalltalk object.

These classes includes: Object, Block, Boolean (True and False), Collection, Array, Dictionary, Interval, OrderedCollection, Date, Error, Exception Warning, Number, String, Time, TimeStamp, TreeModel and so on.

The following image depict how U-Fabrik objects are connected to *real Smalltalk objects*:

In the U-Fabrik Environment instances of a UML Class diagram are instances of Smalltalk class UMLInstance and its class is an instance of UMLClass:

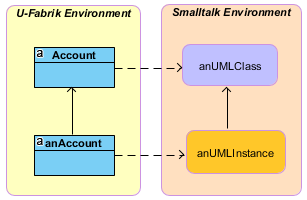


Image 3 - UML Objects representation in both environments

In order to implement behavior the VM has a way to handle message dispatching. The message dispatching is done in the following order:

1. Lookup the method in the class of the receiver object.
2. Lookup the method in all super classes until ‘UndefinedObject’ class is reached.
3. Lookup the method in Traits of the class.

If a method is successfully located then the VM get an instance of UMLParsedImplementation and this instance holds the AST of the method. If not a Debugger with a message does not understand is opened.

Once the method is located then an execution context is created and the VM takes the nodes of the AST and process them one by one. If node is message send then another execution context will be created (for more details see code execution chapter).

## Virtual Machine - Bootstrap

The initial code for basic VM classes is stored as a XML string. The Smalltalk class UMLVirtualMachineBootstrap read the XML string and create the Smalltalk objects that compose the U-Fabrik Virtual Machine.

The bootstrap process create instances of UMLVirtualMachineMetaclass, UMLVirtualMachineClass and UMLVirtualMachineClass and its methods. This process also set the underline Smalltalk class of each U-Fabrik class. U-Fabrik Object class has Object as its underline Smalltalk class, Array has Array and so on. At this point there is hierarchy of U-Fabrik classes with its methods.

Then 3 primitives objects are created **true**, **false** and **nil** as instance of its U-Fabrik classes. From U-Fabrik point of view **nil** is an instance of U-Fabrik class “UndefinedObject” but from Smalltalk point of view is an instance of UMLVirtualMachineInstance. The same applies to **true** which from U-Fabrik point of view is an instance of U-Fabrik class “True” but from Smalltalk point of view is an instance of UMLVirtualMachineInstance.

At the end of the bootstrap process a Smalltalk instance of UMLVirtualMachine with primitive objects (true, false, nil) and basic classes and its methods has been created (the U-Fabrik Virtual Machine is a Smalltalk object).

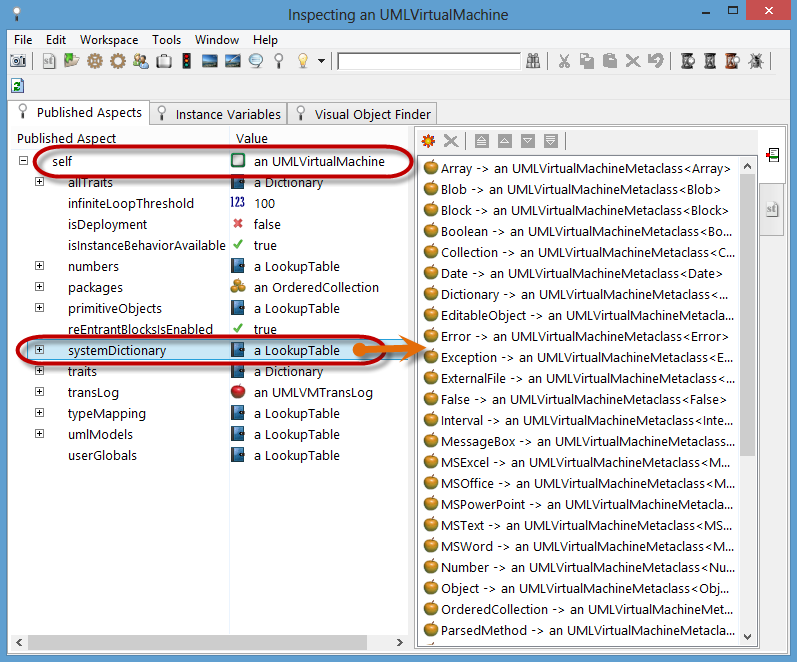


Image 4 - U-Fabrik Virtual Machine object on an inspector

For now this is a one way process so new VM classes and methods has to be added to the XML source in order to be ‘compiled’ in the bootstrap process. Here there is room for improvement to generate the XML source from a running Virtual Machine (see issue [#155](https://github.com/brunobuzzi/U-Fabrik/issues/155)).

## Virtual Machine - Primitives

U-Fabrik Virtual Machine primitives are pure Smalltalk and interact with Smalltalk objects in order to setup U-Fabrik objects. Generally speaking a primitive take a U-Fabrik object (in some cases extract the Smalltalk object such as anArray) and perform some operation that update the state of the objects in the execution context. The operation can be performed in U-Fabrik or Smalltalk realm depending on the primitive.

U-Fabrik Virtual Machine primitives is where Smalltalk and U-Fabrik environment meet each other.

Some Virtual Machine primitives take advantage of the parallelism between both environments and are used to perform different behavior such as <primitive: 101>. The method source in Object class for ~~, ~=, =, == it is the same a call to <primitive: 101>. This primitive is used extensively through the system. OrderedCollection methods at: and asArray are also a call to <primitive: 101>. This primitive basically extract the Smalltalk object from the U-Fabrik object and extract the method name and method arguments from the execution context and execute that in Smalltalk environment. In the final step convert the result to U-Fabrik objects.

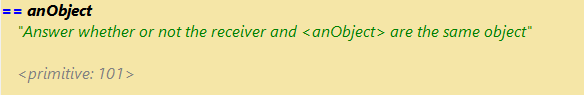


Image 5 - U-Fabrik implementation of == in Object class

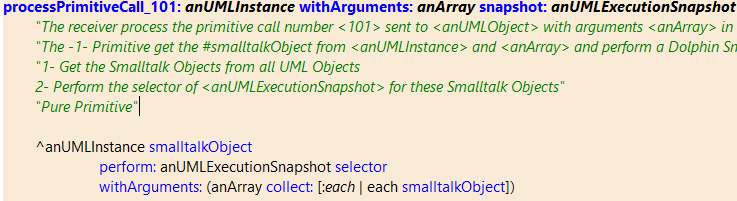


Image 6 - Implementation of primitive 101 at Smalltalk level

Today there are around 96 primitives in U-Fabrik and they are the glue between U-Fabrik and Smalltalk environment.

Other primitives can be easily added to U-Fabrik and incorporate Smalltalk behavior into U-Fabrik is very easily achieved. For example the VM Class TreeModel (in U-Fabrik environment) is a partial implementation of Smalltalk class TreeModel. Most of methods source are a call to primitives 101 or 102. Primitive 102 is very similar to 101 but it does not extract the Smalltalk object from the arguments of the method.

## Virtual Machine - Code Execution

The U-Fabrik Virtual Machine perform code execution using AST (abstract syntax tree) nodes of each method. For each new method source created the UVM creates an instance of UMLParsedImplementation where the AST is stored (as an instance of StMethodNode). The VM know how to process each possible node of an AST. At Smalltalk level all subclasses of StProgram node implement the method processInsideSnapshot:context:parameters: where the first parameter is the execution context (UMLImplementationExecutionSnapshot) the second is the U-Fabrik receiver object (UMLInstance) and the third is an array with method arguments.

The following image show simulation code and the generated AST. In red are highlighted all message send that eventually will create a new execution context.

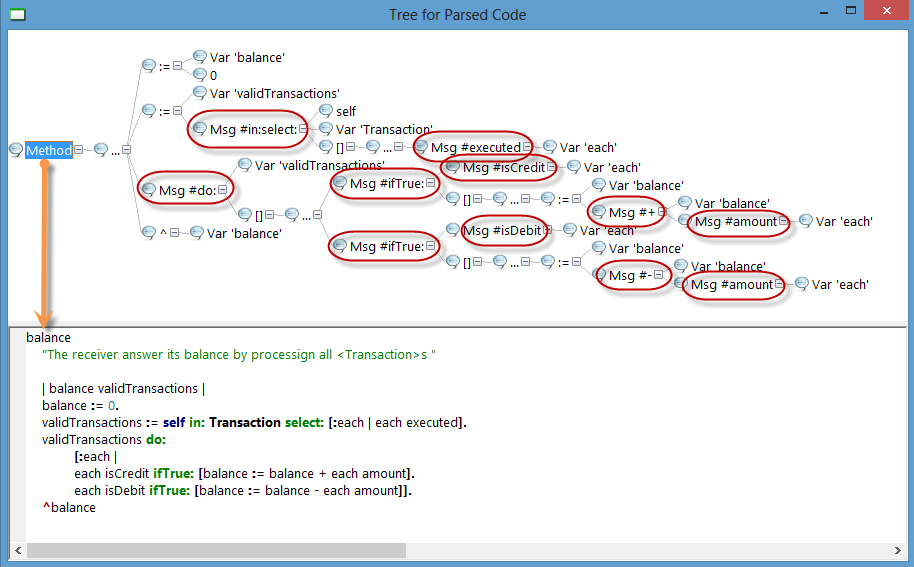


Image 7 - AST (abstract syntax tree) view

When a message is sent to an object the following process is executed:

1. Fetch the implementation based on the receiver and the message. The implementation is an instance of (UMLParsedImplementation).
2. Apply the implementation to the receiver:
   1. Create the execution context (an instance of UMLImplementationExecutionSnapshot).
   2. Iterate over each node of the AST and process it:
      1. Update context objects (reachable objects) with new values.
      2. If there is a message send then this process start again in 1).

The below image depict this process approximately, first the VM fetch the implementation and second apply the implementation to the receiver object. Apply the implementation means to process each node of the AST. Some nodes will trigger another message send (restart the process again in 1 but with a new execution context).

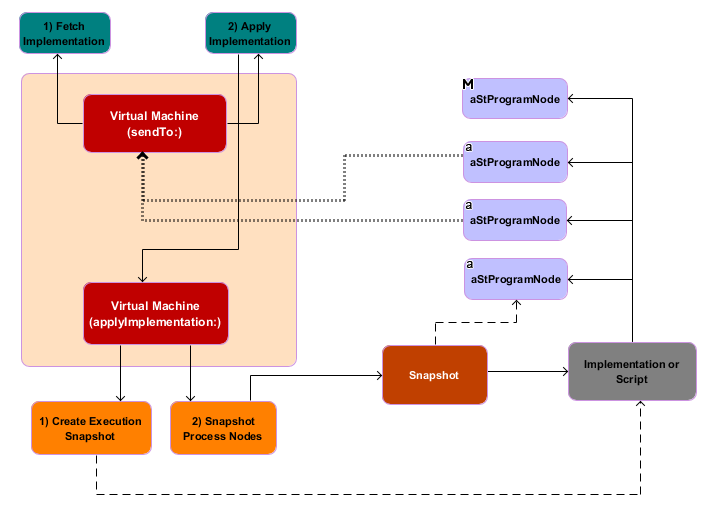


Image 8 - VM AST processing

In each method execution the VM update the state of reachable objects within the execution context or create new objects.

## Virtual Machine - Traits

There are two different form to add behavior to U-Fabrik objects:

1. Adding methods to classes or metaclasses.
2. Creating new Traits.

Traits are composable unit of behavior that are not attached to any class. Only have required methods to be implemented for classes that use them.

The most important Trait in U-Fabrik is “UML Relations Methods”. This trait implement all methods that are used to related objects based on a UML Class diagram. It has 3 required message for the classes:

* getCollectionToOperateWith:
* getCollectionToIterateOver:
* setDirectRelationWith:

All methods are implemented at Object class as primitives so any class can use this Trait.

The primitives that implement these message use the UML Meta Model to establish what kind of relation two object can have based on restrictions in the UML Class diagram.

## Virtual Machine - Blocks

U-Fabrik implement Closures as instance of Smalltalk class UMLBlockClosureNode. Blocks are used as building blocks to create complex simulations of the UML Class diagram.

Each UMLBlockClosureNode has an instance variable that point to the execution context that created the block. Through this execution context all variables used inside the block can be reached.

There are eight primitives to handle blocks: 202 (value), 203 (value:), 204 (valueWithArguments:), 205 (whileTrue:), 206 (whileTrue), 207 (whileFalse:), 208 (whileFalse) and 209 (on:do:). For now U-Fabrik blocks has much less functionality than his Smalltalk counterparts.

In reality all VM classes has the goal to be building blocks to create complex simulations starting with a UML Class diagram. But Blocks are also used to create the basic behavior of the VM such as Boolean implementation, the use of Exceptions and Traits implementation.

The last version U-Fabrik support re-entrant blocks and this was achieved by switching the execution context (UMLImplementationExecutionSnapshot) between each iteration of the Block.

## Virtual Machine – Code File out

All code written for a web a simulation (for UML classes) can be filed out to a XML file and filed in later. This process does not control correctness of the previous code with the new model. If there are compilation failures or other error has to be fixed by hand.

# environment

The U-Fabrik environment has several tools similar (or equal) to any Smalltalk implementation such as:

## Virtual Machine browser

This tool enable user to add new behavior to the virtual machine. Although new primitives can has to be coded in Dolphin Smalltalk access to existing Dolphin methods can be achieved calling primitive 101 or 102.

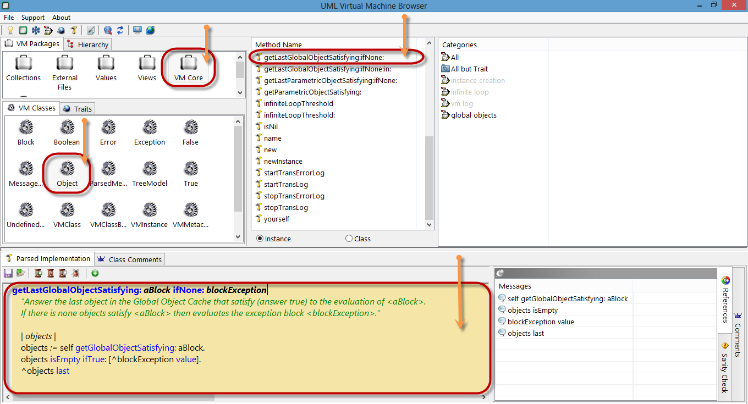


Image 9 - Virtual Machine Browser

## UML Model browser

Installed UML Class diagram are listed in this tool with its classes and methods. The customization of the web application is done here by drag and dropping new method into “Virtual Categories”.

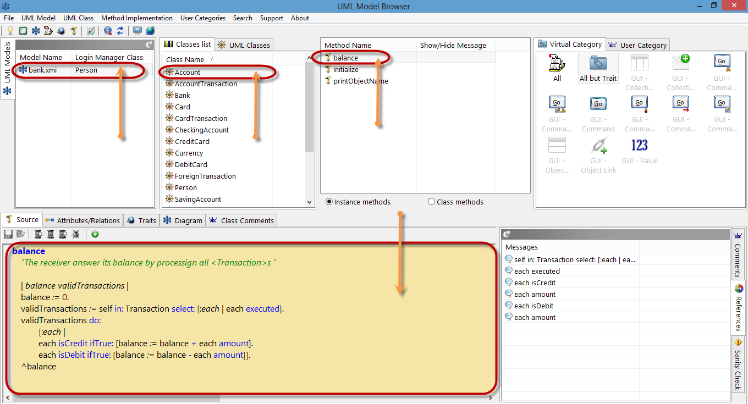


Image 10 - UML Model Browser

### UML Class Attributes and Relations

Display the list of attributes and relations of the current selected class.

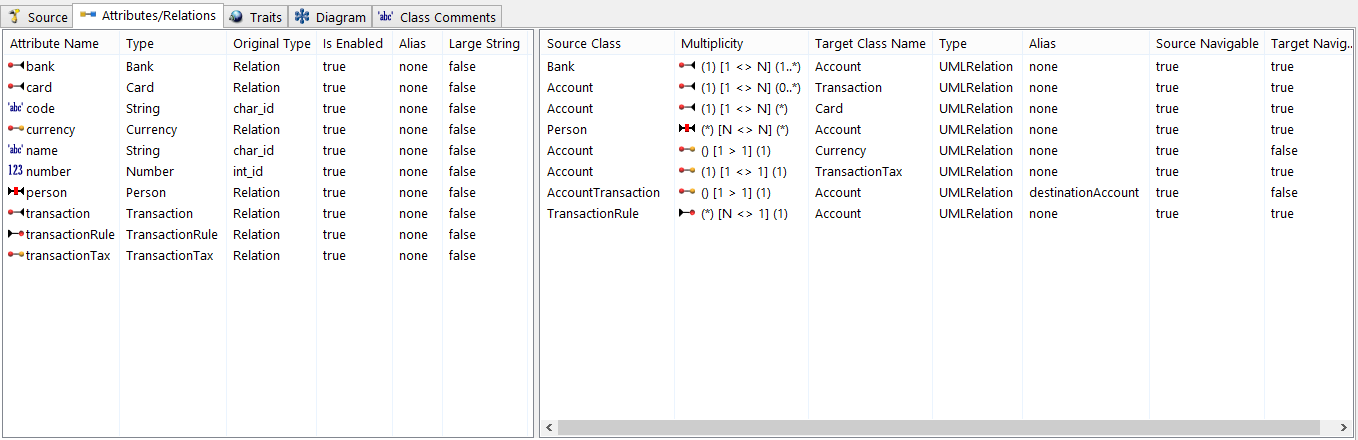


Image 11 - UML Class Relations

### Quick class relations view

A quick graphical view of UML relations of selected class.

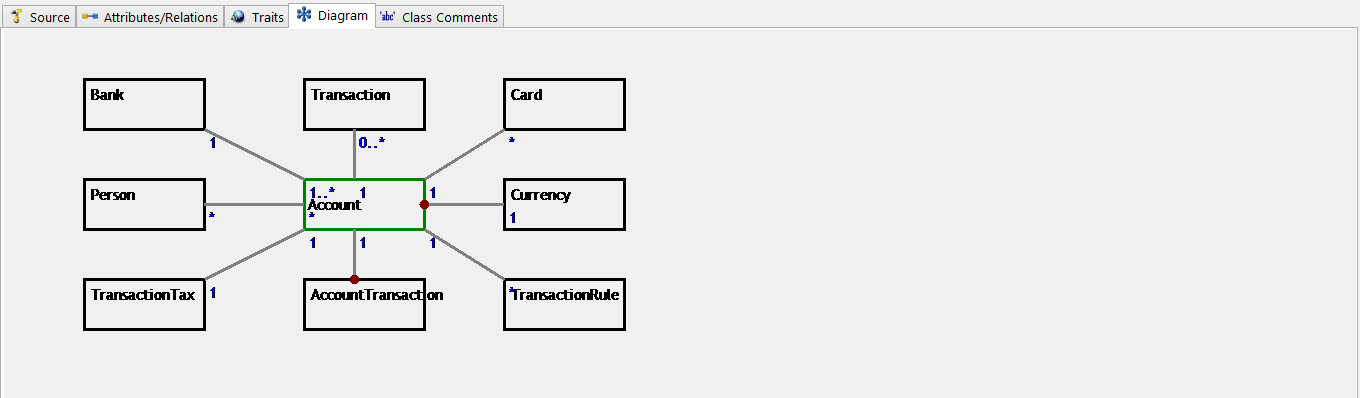


Image 12 - Class relations quick view

## Trait browser

A tool to create new traits that add new orthogonal behavior to objects.

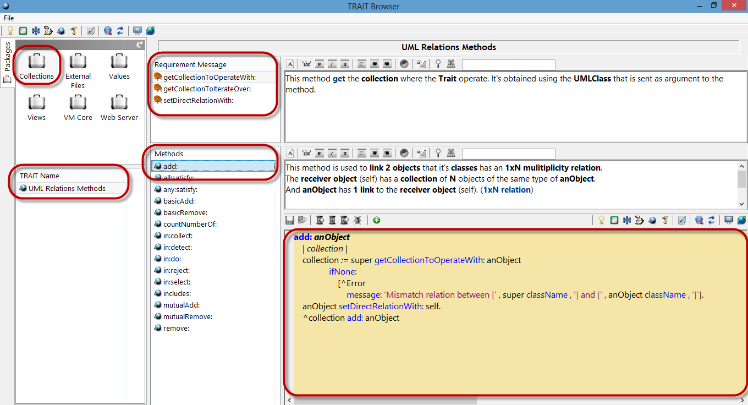


Image 13 - Traits Browser

## UML Script (workspaces)

Classical workspace on U-Fabrik environment to run code scripts.

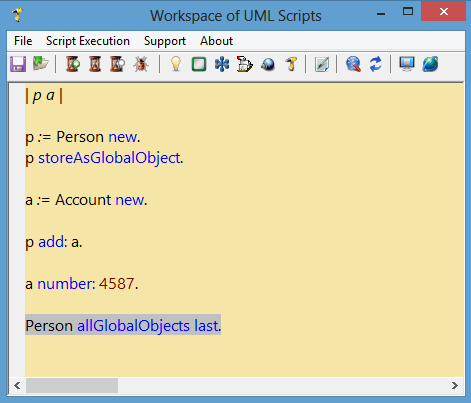


Image 14 - Scripts / Workspace

## Inspectors

A tool explore a UML object and its attributes and related objects.

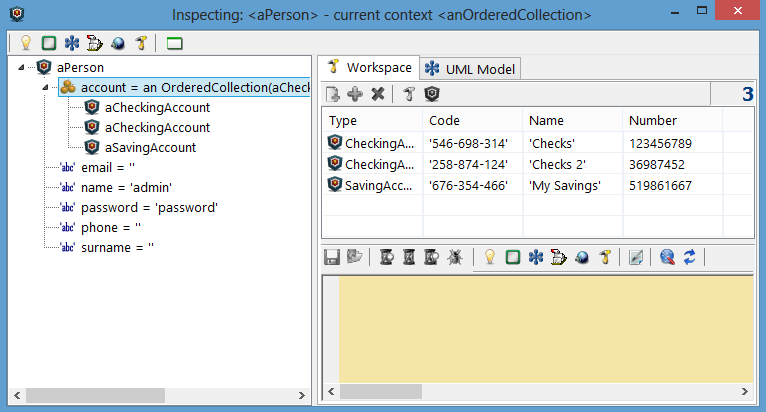


Image 15 - Inspectors

## Debugger

The debugger helps to understand how the code is executed and explore related context objects. It is very basic a needs a lot of improvements in order to have similar capabilities to Smalltalk debugger.

The implementation of the debugger is made with a parallel hierarchy of UMLExecutionSnapshot called UMLDebuggerSnapshot. Maybe instead of further develop the current design of the Debugger another approach is needed.

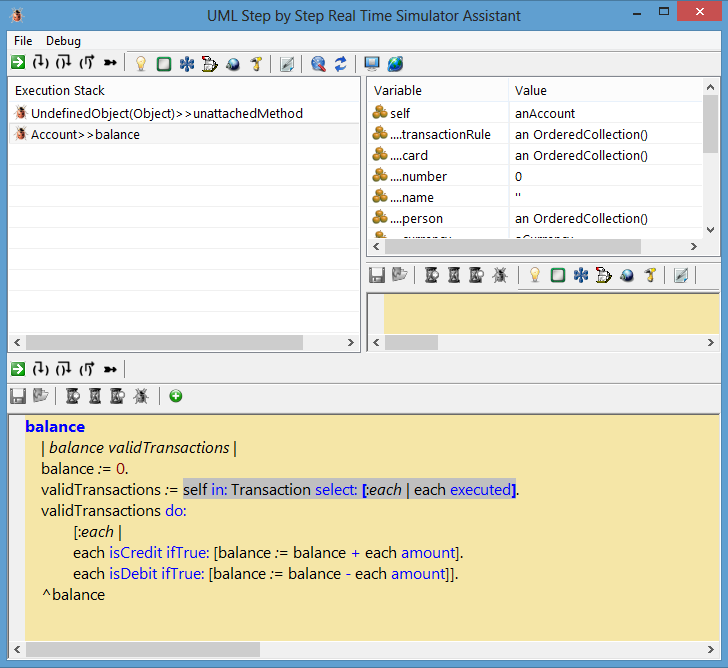


Image 16 - Debugger

## Virtual Machine log viewer

This tool is very useful and display all message sent by the Virtual Machine. It can be enable or disable and also register messages can be deleted.

At Virtual Machine development time this tool can be very useful. It was extensively used implementing re-entrant blocks. The tool allow the inspection of context object (U-Fabrik objects) but also Smalltalk objects that implement the execution context such as UMLExecutionSnapshot and UMLImplementationExecutionSnapshot.

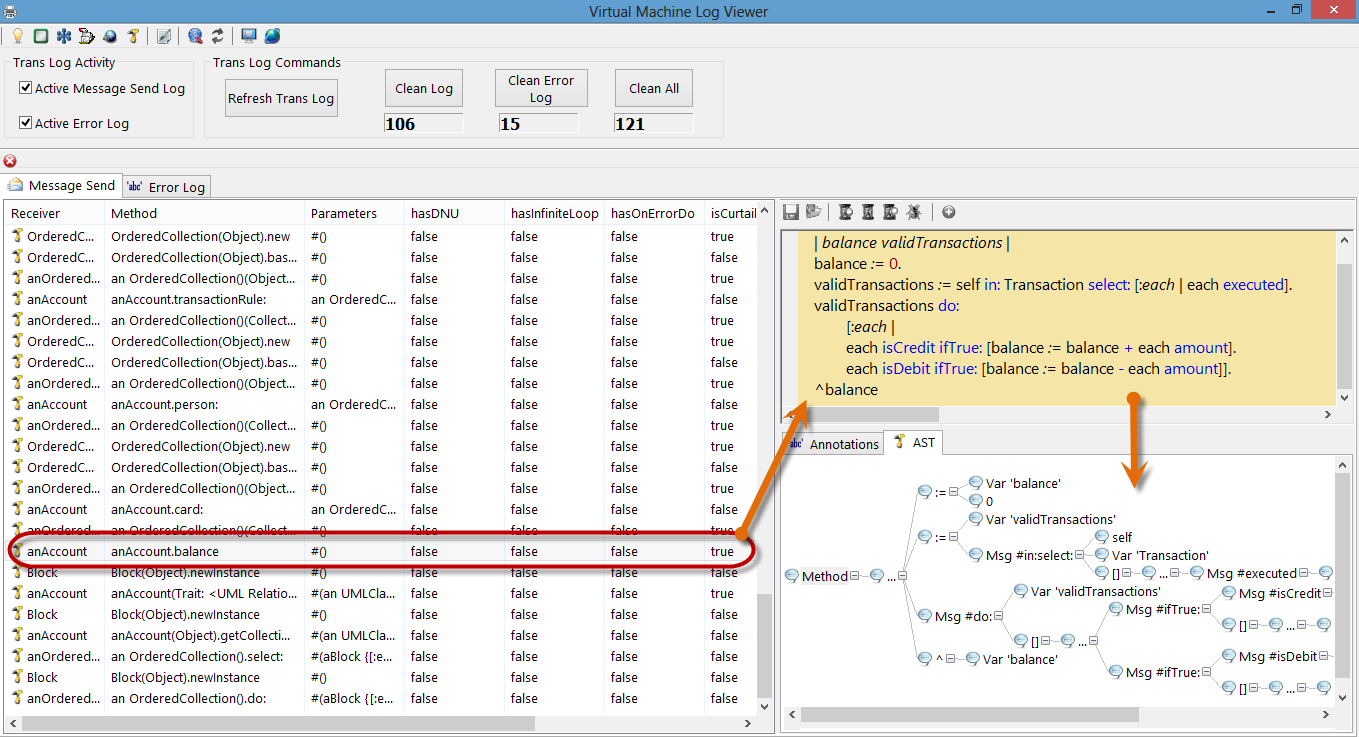


Image 17 - VM log viewer

# web simulation

All previous components of U-Fabrik environment are used to create a simulation of a UML application defined by a UML class diagram.

By default when a UML class diagram is installed in U-Fabrik a standard web simulation is already available for the user. Where each object has its own web page where:

* Simple attributes are rendered directly as simple components (strings, dates, integers, time, booleans, time stamps).
* One to many and one to one relations are rendered as web links to the related object.
* Many to one and many to many relations are rendered as web list (in a separate tab).

The standard web simulation can be further customized by implementing new methods and attach these methods (by drag & drop) to special categories named “Virtual Categories”. There are eleven “Virtual Categories” to customize the standard web simulation.

## Virtual Categories

### **Collection Aspects** (it must answer simple values)

It will render the result of method invocation as object aspects when the object is listed as a web list (each method will be a column).

### **Collection Tab** (it must answer a collection)

It will render the result of method invocation as new web tab with a collection of objects.

### **Command Collection** (can answer any object)

It will render the result of method invocation as web button and it will have a collection of objects as arguments of the invocation. For example set some value and this value is selected from a collection.

### **Command User & Collection** (can answer any object)

The same as the previous but also pass the current user of the web application as an argument.

### **Command** (can answer any object)

The same as two previous but is a web button that trigger a method without arguments.

### **Command Arg** (can answer any object)

It will the result of the method invocation as a web button with arguments. The arguments can be configured.

### **Command External** (can answer any object)

Used to perform a redirect to other url.

### **Command User** (can answer any object)

It will render the result of method invocation as web button and it will have the current user of the web application as method argument.

### **Object Header** (it must answer simple values)

It will rendered the result of the methods invocation as the object header in the web page.

### **Object Link** (it must answer an UML object)

It will rendered the result of the method invocation as object link to other UML object.

### **Value** (it must answer simple values)

It will rendered the result of the method invocation as read only simple web components.

If the method invocation does no answer the expected type then a web text is displayed describing the error.

If any method does not answer the expect object type then a red color text will be rendered displaying the error type.

## Show Hide Method

All methods that render web components can be set (in the UML Browser) with a “*show hide method*”. If this method is not set the web component is rendered directly but if the “*show hide method*” is set then is executed (passing the current web user as argument) before execute the rendering method. If answer **true** then the rendering method is executed after it but if answer **false** the rendering method is not executed and the web component is no rendered. The “*show hide method*” is one argument method and must not fail and answer a Boolean otherwise **false** result will be assumed and the web component won’t be rendered.

In the image below there is a myBanks method that render a tab with a collection of Banks. But tab MyBanks will be renderer only if isAdmin: method answer **true** at rendering time.

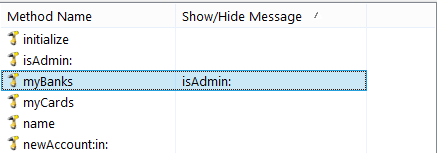


Image 18 - isAdmin: show hide method

The image show the U-Fabrik code for isAdmin: method if the current user has any role that is an *admin* then the web tab **MyBanks** will be rendered.

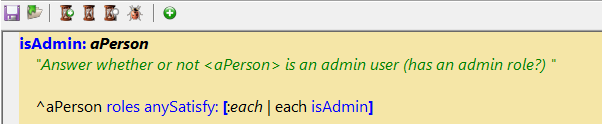


Image 19 - U-Fabrik code for isAdmin:

# Conclusion

This paper highlighted a problem in a specific project context and described the built solution.

First a very specific problem can be solved with a very specific tool but it is also possible to use U-Fabrik in other UML projects.

Second it show the power and flexibility of the Smalltalk environment (in this case Dolphin). It would have been almost impossible to develop U-Fabrik with other development environment (such as Java or C#) in the required time frame and with the listed functionality.

# Future work

There are various places where improvements can be done to the current state. The most important can be:

* **Round trip engineering** **for changes in U-Fabrik to be reflected on the UML case tool**. Right now changes in U-Fabrik are not reflected in the XMI file to be imported in the UML Case tool. This was not a requirement when the system was build.
* **Round trip engineering for VM bootstrap.** Is not possible to generate a XML source of the VM from a running VM. Add classes or behavior to the VM has to be done directly in the XML source.
* **Improvement or re-design** of the debugger.
* **Migrate the web simulation to Seaside.** The web application was developed at the top of AIDA web framework. Seaside has a larger development community.
* Design a better schema to file out UML related code.
* **Simulate inline cache and polymorphic inline cache in method dispatching**. Instead of performing a lookup for a method store the previous methods lookup at call site.

Acknowledgment

The author would like to thank to all Smalltalk communites and especially to Dolphin community and FAST (Argentinian Smalltalk Foundation).

References

1. [↑](#footnote-ref-1)
2. If the UML State diagram is defined as classes in the UML Class diagram then the UML Virtual Machine can simulate the State Machine. For more details: <https://github.com/brunobuzzi/U-Fabrik/wiki/State-Machine-Import> [↑](#footnote-ref-2)