PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/66905

Please be advised that this information was generated on 2018-04-21 and may be subject to change.
Validating Specifications for Model-Based Testing

Submitted to SERP 2008

Pieter Koopman
Software Technology
Radboud University Nijmegen
The Netherlands
Phone/Fax: +31 24 3652(483/525)
Email: pieter@cs.ru.nl

Peter Achten (contact)
Software Technology
Radboud University Nijmegen
The Netherlands
Phone/Fax: +31 24 3652(483/525)
Email: P.Achten@cs.ru.nl

Rinus Plasmeijer
Software Technology
Radboud University Nijmegen
The Netherlands
Phone/Fax: +31 24 3652(644/525)
Email: rinus@cs.ru.nl

Keywords: functional programming, model-based testing, validation tools, quality of specifications

Abstract—In model-based testing the behavior of a system under test is compared automatically with the behavior of a model. A significant fraction of issues found in testing appear to be caused by mistakes in the model. In order to ensure that it prescribes the desired behavior, it has to be validated by a human. In this work we describe a tool, esmViz, to support this validation. Models are given in a pure, lazy functional programming language, esmViz provides an interactive simulator of the model, as well as diagrams of the observed behavior. The tool is built on the iTask toolkit which results in an extremely concise GUI definition. Experiments show that esmViz helps to gain understanding of a model and to detect and remedy errors.

I. INTRODUCTION

In model-based testing the behavior of a system under test, sut, is compared automatically with the behavior of its specification. Examples of model-based test tools are G\/st [5], QuickCheck [2], TorX [9], T-Uppaal [7]. The specification is a possibly nondeterministic state transition system used as model in the tests. The number of states, inputs and outputs can be infinite. The sut is assumed to be a state transition system with a hidden state. One can only apply inputs to the system and observe the corresponding output. Key advantages of model-based test tools are the significant reduction of the amount of manual testing; increase of test speed due to automation; and reuse of specifications for regression testing.

Model-based test systems execute a finite number of traces. For each trace the sut and the specification start in their initial state. An input is selected that is covered by the specification, it is applied to the sut, and the allowed states of the specification are computed. If, during this process, the test system discovers that no states are reachable for the specification, then the sut has shown behavior that is not covered by the specification. In test jargon it is said that an issue is found.

Ideally, each issue indicates an error in the sut. However, in practice a significant fraction of issues appear to be caused by problems with the specification: it does not correctly capture the intentions of the users and the sut does something different. Even though the fraction of issues depends on a lot on factors such as the kind of system and the effort spent in creating the model, we estimate that the specification has to be blamed for about 25% of the issues.

Incorrect specifications are a problem for several reasons. First, if an issue is found it is not clear whether we have to blame the specification or the sut. Finding and correcting errors in the specification takes time during the test phase of the project. Second, errors in the specification are only found during model based testing if the behavior of the sut differs from the specified behavior. Third, any change in the specification during the testing phase can cause major implementation changes to the sut. Finally, any change in model or sut invalidates in principle all previous test results. Hence, errors in the specification can be very expensive and it is worthwhile to invest effort to ensure its quality.

In the model-based test system G\/st the pure, lazy functional language Clean serves as specification language. Due to its high abstraction level it is possible to write concise specifications which contributes to their quality. It allows the test engineer to model arbitrarily large state, input, and output domains exactly as desired. In this work we introduce the tool esmViz to support validation of G\/st models. This simulator enables the user to execute the specification. Such an interactive execution appears to be more illustrative than reviewing the specification. Second, it is possible to record the traces of the specification executed in the simulator. The states visited and their transitions can be visualized in an expanded state transition diagram. Since the type of states, inputs and outputs can be infinite and different in each and every specification, doing this conveniently is not straightforward. The key to the solution is to use generic definitions such that operations on these types can be derived instead of defined manually.

The layout of the paper is as follows: in Sect. II we introduce the concepts and notation that will be used throughout this paper. In Sect. III we discuss the issues that arise when testing against a formal specification. In Sect. IV we describe esmViz. Its implementation is discussed in Sect. V. Related work is discussed in Sect. VI. We present user experiences in Sect. VII, and conclude in Sect. VIII.

II. MODEL-BASED TESTING

In model-based testing the test tool compares the observed behavior of the system under test, sut, with the model in order to judge the correctness of the behavior. Any deviation of the observed behavior of the sut from the behavior allowed by the model is called an issue. In this section we review the models used by the model-based test tool G\/st.

The models used by G\/st for testing state based systems are extended state systems, ESMs. An ESM consists of some initial state
A. Conformance

possible target states \( t \), reachable after the given trace \( a \):

B. Testing Conformance

the model by executing a finite number of traces of finite length.

If we have a trace \( s \Rightarrow t \) and a transition \( t \xrightarrow{i/o} u \) we can construct the trace \( s \xrightarrow{i/o} u \). If we are not interested in the

target state, we will occasionally write \( s \xrightarrow{i/o} \) \( \equiv \exists t \cdot s \xrightarrow{i/o} t \) and \( s \xrightarrow{i/o} \equiv \exists t \cdot s \xrightarrow{i/o} t \).

All traces from a given state are defined as: 

\[
\text{traces}(s) \equiv \{ \sigma \mid s \xrightarrow{\sigma} t \}
\]

The init of a state \( s \) is the set of inputs \( i \), such that there is an output \( o \) and target state \( t \) in the ESM such that

\[
\text{init}(s) \equiv \{ \sigma \mid s \xrightarrow{\sigma} t \}
\]

or a set of transitions of the form \( s \xrightarrow{i/o} t \). The after of a state \( s \) is the set of possible target states \( t \), reachable after the given trace \( \sigma : s \xrightarrow{\sigma} t \). We overload traces, init, and after for sets of states

instead of a single state by taking the union of the set members.

A. Conformance

In model-based testing we try to determine conformance of the

sut and the model called spec. The sut is assumed to be a transition system, but treated as a black box: one can observe its traces, but not its internal state. During tests, all observed traces of the sut have to be traces of the specification to say that the sut conforms to the specification. Formally, this relation is defined as:

\[
\text{sut} \conform \text{spec} \equiv \forall \sigma \in \text{traces}_{\text{spec}}(s_0),
\]

\[
\forall i \in \text{init}(s_0), \text{after}_{\text{spec}}(i), \forall o \in O,
\]

\[
(t_0, \text{after}_{\text{sut}}(i) \xrightarrow{i/o} (s_0 \text{ after}_{\text{spec}}(i)) \xrightarrow{i/o}
\]

Here \( s_0 \) is the initial state of spec, and \( t_0 \) the initial state of sut. Intuitively the conformance relation reads: if the specification allows input \( i \) after trace \( \sigma \), then the observed output of the sut should be allowed by the specification. If spec does not specify a transition for the current state and input, anything is allowed. Because the sut is a black box, its initial state \( t_0 \) is generally not known explicitly. We assume that the sut is in this abstract state when we switch it on, or we reset it.

Limiting the applied inputs to the init of the states of the current traces allows for partial specifications spec.

B. Testing Conformance

The conformance relation defined above covers all traces. Most interesting systems contain cycles, so traces can become infinitely long. Due to the possible infinite types for input and output, there can be even infinitely many traces of finite length. It is clear that in general a test system cannot prove conformance by executing tests. The test system \( \text{G\text{\textcircled{\textit{v}}}s\text{\textit{t}}} \) approximates the conformance of the sut to the model by executing a finite number of traces of finite length.

To increase efficiency the test system records the set of allowed states, \( s_0 \text{ after } \sigma \), rather than the trace \( \sigma \). If at some point in the test this set of states becomes empty we have found an issue: a trace that shows that there is no conformance between sut and the model. Clearly this way of testing is sound, each trace leading to an issue during testing shows that there is no conformance between the sut and the model. This way of model-based testing is also complete, if there is no conformance between sut and the model, there are one or more traces indicating this. Such a trace can be found by testing (if the allowed length during testing is sufficiently large).

C. Representation of the transitions

To represent the ESM in the model-based test tool \( \text{G\text{\textcircled{\textit{v}}}s\text{\textit{t}}} \) we need a finite (preferably small) and flexible representation, even if the set of transitions is infinite. Furthermore it should be easy to determine the init of the set of actual states, or to determine if an input is in this set, since this information is needed before we can apply an input during model based testing. The crucial step is to use a function to model the transitions rather than a data structure containing individual transitions. Each function alternative with variables in its patterns captures a family of related transitions. As usual lists represent sets. To define init easily we use specifications of type \( S \times I \rightarrow \left[ \text{Trans O S} \right] \).

A basic assumption in \( \text{G\text{\textcircled{\textit{v}}}s\text{\textit{t}}} \) is that a transition always contains a sequence (list) of output symbols. This gives some additional flexibility as well as a suitable notation for no output (the empty list). Usually it is most convenient to specify the sequence of outputs and the target state in a transition. However, the number of allowed output sequences for one input can get huge, which makes it infeasible to state them explicitly. For instance in an authentication procedure a typical step is to ask for a challenge (the input), the response is a 64 bit number. Listing all possible outputs and target states explicitly requires \( 2^{64} \) transitions. In such a situation we prefer one function of type \( \left[ \text{O} \right] 

This specification is partial (e.g. the effect of pressing a product button when the machine in the state Off is not defined), and nondeterministic (if there is enough money in the machine and the user asks for coffee, the machine either produces coffee, or does nothing at all). Non-determinism models limited knowledge of the state of the real machine: e.g. if there are coffee beans it will produce coffee, otherwise it cannot produce coffee and waits for a new command.

III. ISSUES FOUND IN MODEL BASED TESTING

Issues are traces that show that there is no conformance between the sut and the specification. Ideally each issue found indicates an error (bug) in the sut, but that is not always the case. Other sources of issues are inaccuracies in the model, problems in the interface between the test system and the sut, and internal faults in the test tool. One wishes to eliminate these other sources of issues before actual testing starts.

In ordinary automatic testing the test tool executes a manually specified or recorded trace. As a rule of thumb test engineers say that 40% of the issues found indicate a real error in the sut. A tiny fraction of these issues is caused by the test tool itself, or the interface with the sut. Most issues are caused by the fact that the trace used does not correspond to the current version of the specification, or the specification itself is incorrect.

In model-based testing the traces are generated automatically and on-the-fly from the specification. This guarantees that the traces used during the tests always correspond to the current specification. As one expects this implies that a larger fraction of the issues found indicate real errors in the sut. A tiny fraction of these issues is caused by the test tool, or the interface with the sut. Most issues are caused by the fact that the trace used does not correspond to the current version of the specification, or the specification itself is incorrect.

In model-based testing the traces are generated automatically and on-the-fly from the specification. This guarantees that the traces used during the tests always correspond to the current specification. As one expects this implies that a larger fraction of the issues found indicate real errors in the sut. A tiny fraction of these issues is caused by the test tool, or the interface with the sut. Most issues are caused by the fact that the trace used does not correspond to the current version of the specification, or the specification itself is incorrect.

In model-based testing the traces are generated automatically and on-the-fly from the specification. This guarantees that the traces used during the tests always correspond to the current specification. As one expects this implies that a larger fraction of the issues found indicate real errors in the sut. A tiny fraction of these issues is caused by the test tool, or the interface with the sut. Most issues are caused by the fact that the trace used does not correspond to the current version of the specification, or the specification itself is incorrect.

The specification is a Clean function, hence the compiler can readily check relevant properties: i) are all used identifiers properly defined, ii) is the entire specification type correct, iii) are all alternatives (transitions) reachable. Still, well typed specifications can go wrong. The problems with specifications that cannot be detected by the compiler can be divided in the following classes.

1) Relevant behavior of the system is not covered in the specification. Since the test system is carefully designed to handle partial specifications, this cannot be detected. Missing parts of the behavior are not covered in the tests.

2) The specification contains design errors. Typically a family of transitions is too large or too small, or leads to the wrong target state. If the sut does a better (or at least different) job, the test system will notice the difference if an appropriate trace occurs and hence report an issue. Consider the alternative for vSpec (On s) (Coin c) in the example of the previous section. The wrong condition and forgotten return of money if the state becomes too large are probably design errors.

3) The transitions are not always defined. A typical example is the use of lowercase identifiers (variables) where an uppercase identifier (constructor) is intended, or vice versa. Another source of problems is copy-paste programming used to define similar transitions, where not all necessary changes are made. In our example this occurs in the transitions for vSpec (On s) (Butt coffee), and vSpec (On s) (Butt Tea).

All these problems result in well typed models. If the implementation is based on such a model, it is not possible to detect the problems by testing. Nevertheless, they must be found and preferably before model-based testing starts.

Various approaches to find these kind of problems are: Inspection or reviews of the specification. Problems can be found by manual inspection of the specification. As the model tells the whole story, there is nothing that prevents these errors from being detected by reviewing the code. However, due to their subtle nature, they might be missed.

Model checking. If we have the right properties and the specification is available in a form suited for model checking, the problems can be found by model checking. Limiting factors are the availability of the model in a form suited for a particular model checker, and the availability of properties to check. If the problems are known we can often find such a set of properties quite easily, but that is too late. In our example we can require: p1) every transition preserves the amount of money, p2) the amount of money in the machine is always less then max, and p3) if we receive a product, it must be equal to the requested product. Finding a complete set of properties that reveals all problems is in general quite tricky.

Testing properties of the specification. Properties on transitions can be tested by the logical branch of GVst. The advantage is that everything can be done within the same framework, especially the Clean specification function can be used as subject of tests. The drawback is that testing gives less certainty for large systems (although for small specifications the logical test system provides a proof by exhaustive testing).

Validation by simulation. The specification can be used as basis for an interactive simulation. With some effort the simulator not only displays the current transition, but also depicts the state space that is covered in the current simulation. Such a simulation can reveal that important parts of the behavior are missing, as well as problems with individual transitions. This requires a thorough observation of the shown behavior. Since the state space is discovered step-by-step by the user, the chances of finding the problems are quite good.

Each of the above methods can in principle find problems in the specification, but none of them can guaranteed this. Each method either requires human spotting of problems, or human formulation of properties revealing the problems.

In the remainder we describe esmViz. It combines model-checking of properties on transitions with validation by step-wise simulation. Together with GVst, this covers a broad range of tools to investigate the quality of models.

IV. VALIDATION OF SPECIFICATIONS WITH esmViz

In this section we describe the web browser-based simulation tool, esmViz, that we have created to determine the quality of specifications. The tool also gives an impression of the behavior specified by the model, and checks user defined predicates on the transitions encountered. Simulation is useful to give non-experts a good impression of the specified behavior. The GUI of esmViz is a screen with the following elements (Fig. 2(a)): 1. A list of found issues. The list is empty in Fig. 2(a). 2. The explored model as an Extended State Diagram (ESD). 3. Within the ESD the set of possible active states determines the inputs that can be given. These are enumerated as buttons that the user can press to advance the system one step. In Fig. 2(a) the active states are S = \{0, 20\}, and init (S) = {ButtCoffee, ButtTea, Coin0, Coin20, SwitchOff}. 4. Commands for navigation purposes, resetting the exploration, and so on. 5. The current trace, as explained in Sect. II. Here the trace
A. The ESM description

The ESM is a formal model of a system's behavior, specified as a directed graph. In ESM diagrams, a parameterized state is drawn as one state. Each state is defined for each value of the parameters encountered during simulation. The ESM is a graphical representation of the system's behavior.

The tool esmViz creates an ESD of the behavior encountered during simulation, which is rendered as a directed graph. In ESM diagrams, a parameterized state is drawn as one state. Each state is defined for each value of the parameters encountered during simulation. The ESM is a graphical representation of the system's behavior.

The ESM is a formal model of a system's behavior, specified as a directed graph. In ESM diagrams, a parameterized state is drawn as one state. Each state is defined for each value of the parameters encountered during simulation. The ESM is a graphical representation of the system's behavior.

B. Example

The ESM is a formal model of a system's behavior, specified as a directed graph. In ESM diagrams, a parameterized state is drawn as one state. Each state is defined for each value of the parameters encountered during simulation. The ESM is a graphical representation of the system's behavior.

The tool esmViz creates an ESD of the behavior encountered during simulation, which is rendered as a directed graph. In ESM diagrams, a parameterized state is drawn as one state. Each state is defined for each value of the parameters encountered during simulation. The ESM is a graphical representation of the system's behavior.

The ESM is a formal model of a system's behavior, specified as a directed graph. In ESM diagrams, a parameterized state is drawn as one state. Each state is defined for each value of the parameters encountered during simulation. The ESM is a graphical representation of the system's behavior.

The tool esmViz creates an ESD of the behavior encountered during simulation, which is rendered as a directed graph. In ESM diagrams, a parameterized state is drawn as one state. Each state is defined for each value of the parameters encountered during simulation. The ESM is a graphical representation of the system's behavior.

The ESM is a formal model of a system's behavior, specified as a directed graph. In ESM diagrams, a parameterized state is drawn as one state. Each state is defined for each value of the parameters encountered during simulation. The ESM is a graphical representation of the system's behavior.
where healthy checks \( p_1 - p_3 \) (Sect. III). An ESD showing all issues described by healthy is depicted in Fig. 2(b).

healthy :: (SeenTrans State Input Output) \( \rightarrow \) [[String]]
healthy (s,i,o,t) = [if \( (vs+vi \neq vo+vt) \) // value preservation in transition? \( (p_1) \)
  ["value is not preserved in this transition, 
  ","value s+value i=", toString (vs+vi)
  ", and value o+value t=",toString (vo+vt)] 
  ,if \( (vt > Max) \) // value of target state within bound? \( (p_2) \)
  ["Value of target state ",toString vt
  ", larger than Max (",toString Max,")." ] ]
  , case (i,o) of // obtained the ordered product? \( (p_3) \)
  (Butt p,[Gxp q]) | p !== q
  = ["The required product is unequal*,
  ", to the delivered product!*]
  _  = []
  ]
where vs = value s; vi = value i
vo = value o; vt = value t

V. IMPLEMENTATION

The esmViz tool has been written in Clean, using the iTask toolkit [8]. Despite its conciseness (800loc) it offers a fair amount of functionality (see also other tools in Sect. VI). In this section we present the most interesting parts of the implementation. These are the main structure of the GUI (Sect. V-A) and the integration of the ESD rendering tool Graphviz [3] that we used in the application (Sect. V-B).

A. The Main GUI Structure: Iterating iTasks

The main GUI structure of esmViz is an iteration of the main tool function DiGraphFlow. As described in Sect. IV, it provides the user with a number of elements, expressed as a list of choices (the arguments of orTaskL below which folds the basic iTasks choice user with a number of elements, expressed as a list of choices (the task function DiGraphFlow. As discussed in Sect. IV, the user can select a new state. For reasons of space, we do not show the code of the other functions.

B. The Rendering of the Exploded Automaton

By far the most intricate component of the GUI is the ESD editor. Creating attractive renderings of directed graphs is known to be a hard problem. Fortunately, we can rely on other tools to solve this problem. Here we have used the Graphviz tool set [3]. Directed graphs are described using the DOT language. Given a DOT text file, the dot tool can be invoked to create a rendering in various formats (we will use the gif output). Note that this interface is text-based, whereas editors in the iTask toolkit are type based. We can embed the text based tools of Graphviz in the type based iTask toolkit in a compositional way by defining a suitable collection of data types that describe an ESD as a directed graph. This collection of data types captures the DOT language. The relevant top level type definitions are:

:: Digraph = Digraph String [GraphAttribute] [NodeDef]
:: NodeDef = NodeDef Int [NodeAttribute] [EdgeDef]
:: EdgeDef ::= (Int, [EdgeAttribute])
:: SelectedItem = Node Int

A (Digraph nameatts nodes item) value represents a directed graph. A directed graph has nodes, each of which is identified by a number, and is connected with other nodes by means of edges. Graphs, nodes, and edges have attributes. Graphviz supports an extensive set of attributes (almost 150) that can be used to alter and tweak the output. In DOT, attributes are specified as name = value pairs. Some attributes are shared by graphs, nodes, and edges. We have represented attributes separately for graphs, nodes, and edges, each as a list of unary data constructors. For instance, for graph attributes we have Get_name value pairs. A single generic function prints these values as correct DOT expressions. The result is that we have both a typed representation of DOT expressions (Digraph values) as well as a textual representation (printing such a value with toString). The function mkDiGraph yields the Digraph value that represents the currently explored ESD.

The iTask editor for Digraph values performs the following actions for a d :: DiGraph value identified by name. First, compute e = toString d and save e in file name.dot. Second, invoke dot on name.dot, which yields a rendering as name.gif. Third, invoke dot to create a name.map file to allow the user to select states. Fourth, alter the lines in name.map to invoke a script that sends the label of the selected state to the server application. Finally, generate the proper HTML to be included in the application page. The server application, when receiving the label of a selected state, updates the corresponding Digraph value to reflect the change. Now the application continues with the new Digraph value.

VI. RELATED WORK

The mCRL2 tool set [7], [7] uses a process algebraic specification language, mCRL2 [7], to describe distributed, communicating systems. It has a functional style data language with recursive types, data constructors, functions, lambda-abstraction, and structured data. It comes with an extensive number of tools (15) for analysis purposes. Five are relevant to our work: with xsim a user can explore a linearized mCRL2 specification in a similar way as with our tool, using a GUI (the simulation tool sim has a command line interface): the user can select actions, after which the tool shows the resulting state. Besides interactively exploring the mCRL2 specification, the tool set also allows to render the complete state space: NoodleView (for 2D rendering) and FSMView (for 3D rendering). Before this is possible, the state space needs to be generated with lps2lts.
The TorX tool set [9], [7] is a model based test tool to check conformance of real suts, based on the ioco theory of testing. The specification is a Labelled Transition System (LTS), or one that is derived from a higher level specification language that converts to LTS (e.g. mCRL2 described above). The tool uses the specification to automatically determine inputs, observe outputs from the sut, and make a final verdict. In this sense, it is not useful for exploring a specification. However, once a test run has been created, the user can explore the actual trace which is depicted as a message sequence chart.

The Uppaal tool set [1], [7] can be used for both validation and verification (using model checking) of time-based systems. Validation is done by means of a graphical simulator of a time-based automaton specification. The automaton specification is basically a labeled transition system with timing constraints. Uppaal allows for simple data types, clocks, and constraints on these clocks. The user can create specifications in an intuitive, graphical way. The user can stepwise direct the system’s behavior, or generate a random trace.

The esmViz tool differs with the mCRL2 approach in that we use a single modeling formalism. Except for the 3D rendering all of the functionality of the mCRL2 tool set is available in esmViz. The TorX tool set is really a model based testing harness, and is less suited for exploration purposes. Specifications within Uppaal can be created graphically. In esmViz specifications are given as a function, out of which a graphical approximation is ‘discovered’ by the user or by the system. In our opinion this combines the best of both worlds: the succinctness of functional programming with the intuitive appeal of a graphical rendering.

VII. EXPERIENCES

In order to judge the quality of esmViz 10 master students in computer science studied some test cases with and without esmViz. These students are literate clean programmers, have a basic understanding of model-based testing with G
cst and the specifications needed (but no hands-on experience). After an introduction to esmViz and playing with an example similar to the beverage vending machine in this paper the students were asked to locate problems in two other case studies. The examples were heavily parameterized specifications of a number guessing game and a telephone number derived from a higher level specification language that converts to rendering them.

The students found esmViz very handy to get a feeling for the behavior of the specified system. Everybody found it much easier to understand a specified system with the tool than without. Finding errors in the specification by simulation remains hard, but the tool makes it easier. The same holds for finding the source of issues found by G

cst. This is consistent with the general observation in all kinds of testing: finding issues is one thing, but finding their cause is another.

VIII. CONCLUSIONS

There are two kinds of conclusions from the work described in this paper. First, the specification simulator esmViz described in this paper really helps a lot to understand the behavior of the extended state machines used as specification in model-based testing. Although the compiler of the statically typed functional programming language used as carrier of these specifications checks the models, the models can still contain errors. Finding these semantic errors is hard. The simulator helps in locating these problems, especially if an appropriate constraint on transitions or states is known. Second, implementing such a tool with iTasks is a real pleasure. Integrating Graphviz with iTasks turned out to be smooth. Implementing a browser interface for esmViz using the iTask system imposes some restrictions on the layout of the GUI, but works well. The different possible user actions are modeled each by an iTask. The iTask system is well suited to compose these tasks in a flexible way and takes care of rendering them.

REFERENCES