| 著者       | 高橋 哲也       |
| 出版者     | 法政大学大学院情報科学研究科 |
| 原作名     | 法政大学大学院紀要　情報科学研究科編 |
| 期間      | 2019-03-31 |
| 発行者     | ホセイ大学                                      |
| ISSN       | 0002-1922                                           |
| URL        | http://doi.org/10.15002/00021922 |

### 摘要

論文の目的は、システムマネジメント（SysML）活動図を対象とした形式的検証手法であるテスティングベースの形式的検証（TBFV）を提案し、その有効性を示すものである。TBFVは、活動図の動作を適切に保証するための手法であり、システムの設計をより効率的に実現するのに役立つことが期待される。

### 方法

TBFVの手法は、活動図の動作をシミュレーションに基づいて検証する。シミュレーションは、活動図の動作を数値データとして記録し、それを基に検証が実行される。

### 結果

TBFVの検証手法は、活動図の動作を適切に保証することができ、システムの設計をより効率的に実現するのに役立つことが示された。

### 考察

TBFVの手法は、活動図の動作を適切に保証するための手法であり、システムの設計をより効率的に実現するのに役立つことが期待される。

### まとめ

本論文は、システムマネジメント（SysML）活動図を対象とした形式的検証手法であるテスティングベースの形式的検証（TBFV）を提案し、その有効性を示すものである。TBFVの手法は、活動図の動作を適切に保証するための手法であり、システムの設計をより効率的に実現するのに役立つことが期待される。
TBFV-M: Testing-Based Formal Verification for SysML Activity Diagrams

Yufei Yin
Graduate School of Computer and Information Sciences
Hosei University
Tokyo 184-8584, Japan
yufei.yin.59@stu.hosei.ac.jp

Abstract—SysML activity diagrams are often used as models for software systems and its correctness is likely to significantly affect the reliability of the implementation. However, how to effectively verify the correctness of SysML diagrams still remains a challenge. In this paper, we propose a testing-based formal verification (TBFV) approach to the verification of SysML diagrams, called TBFV-M, by creatively applying the existing TBFV approach for code verification. We describe the principle of TBFV-M and present a case study to demonstrate its feasibility and usability. Finally, we conclude the paper and point out future research directions.

Keywords—SysML activity diagrams; TBFV; test path generation; formal verification of SysML diagram;

I. INTRODUCTION

Model-Based Systems Engineering (MBSE) [1] is often applied to develop large scale software systems in order to effectively ensure their reliability and to reduce the cost for testing and verification. The systems modelling language SysML [2, 3] can support effective use of MBSE due to its well-designed mechanism for creating object-oriented models that incorporate not only software, but also people, material and other physical resources. In MBSE, SysML models are often used as the design for code. Therefore, its correctness in terms of meeting the users’ requirements becomes critical to ensure the high reliability of the code. Unfortunately, to the best of our knowledge from the literature, there are few tools to support the verification of SysML models [4, 5] in particular rigorous ways of verification.

Testing-Based Formal Verification (TBFV) proposed by Liu [6-8] shows a rigorous, systematic, and effective technique for the verification and validation of code. Its primary characteristic is the integration of the specification-based testing approach and Hoare logic for correctness proof of code to guarantee the correctness of all the traversed program paths during testing. The advantage of TBFV is its potential and capability of achieving full automation for verification through testing. However, the current TBFV is mainly designed for sequential code in which all of the details are formally expressed, and there is no research on applying it to verify SysML models yet. In this paper, we discuss how the existing TBFV can be applied to SysML models for their verification and we use TBFV-M (testing-based formal verification for models) to represent the newly developed approach. Since SysML Activity Diagrams can model the systems dynamic behavior and describe complex control and parallel activities, which are similar to code but with additional constructs such as parallel execution, our discussion in this paper focuses on the activity diagrams.

The essential idea of TBFV-M is as follows. All of the functional scenarios are first extracted from a given formal specification defining the users’ requirements where each functional scenario defines a meaningful functional behavior of the system. Meanwhile, test paths are generated from corresponding SysML Activity Diagrams waiting to be verified. Then, test paths are matched with functional scenarios by comparing the collection of decision condition of each test path and the guard condition of the functional scenario. After this, the pre-condition of the test path is automatically derived by applying the assignment axiom in Hoare logic based on the functional scenario. Finally, the implication of the pre-condition of the specification in conjunction with the guard condition of the functional scenario to the derived pre-condition of the path is verified through automatic proof or testing to determine whether the path contains bugs. The details of this approach will be discussed from Section 5.

The remainder of the article will detail the TBFV-M method. Section 2 presents related work we have referenced. Section 3 introduces the Testing-Based Formal Verification technique for the verification and validation of code. Section 4 mainly details the whole development process of using Model-Based Systems Engineering and the application scenarios of TBFV-M method. Section 5 describes the principle of TBFV-M, showing the core technology of TBFV-M. Section 6 uses one case study to present the key point of TBFV-M. Finally, the details of the implementation of the algorithm are presented in Section 7. Section 8 characterizes the evaluation part. Section 9 concludes the paper.

II. RELATED WORK

In this section, we briefly review the existing work related to our study. For the sake of space, we focus on those we have referenced during our research. We divide the related work into four different parts, including testing-based verification, requirements verification, verification using Hoare Logic and test case generation.

Considering the shortcoming of formal verification based on Hoare logic being hard to automate, Liu proposed the TBFV (Testing-Based Formal Verification) method by combining specification-based testing with formal verification [6]. This
method not only take the advantage of full automation for testing, but also the efficiency of error detection with formal verification. Liu also designed a group of algorithms [9] for test cases generation from formal specification written in SOFL [10]. A supporting tool [8] is also developed. These efforts have significantly improved the applicability of formal verification in industrial settings.

Franco Raimondi [11] addressed the problem of verifying planning domains written in the Planning Domain Definition Language (PDDL). First, he translated test cases into planning goals, then verified planning domains using the planner. A tool PDVer is also generated. In this paper, testing is also used during verification and the effectiveness and the usability is improved.

Stefano Marrone [12,13] designed a Model-Driven Engineering approach, in which formal models are constructed and test cases are generated from UML model, utilizing UML profiles and model transformation algorithms, automatically. As they claimed, formal models can be used for quantitative analysis of non-functional properties, while test cases can be used for model checking. A railway signaling example is shown to introduce its integration, usability and reduction of manual activities.

Feng Liang [14] proposed a vVDR (Virtual Verification of Designs against Requirements) approach for verifying a system with its requirement. In his research, the system is modeled in Modelica, and requirement verification scenarios are specified in ModelicaML, an UML profile and a language extension for Modelica. vVDR approach guarantees that all requirements can be verified by running this scenario automatically. However, the deficiency appears when the number of requirements and scenarios increase.

Inspired by Liu’s work, we apply and extend the TBFV approach to models and propose the TBFV-M. A model is more intuitive than a formal specification because it requires less relevant background knowledge and is easier to communicate with customers. TBFV approach shows the treatment of code, while TBFV-M approach deals with SysML Activity Diagrams. And different with Feng Liang’s work, TBFV-M approach do not use other supporting tools, like Modelica, we merely use Hoare Logic to do the verification. Referring to test case generation, TBFV-M approach can deal with unstructured diagrams, which may have stronger processing power than existing approaches.

III. INTRODUCTION OF TBFV FOR CODE

TBFV is a novel technique that makes good use of Hoare logic to strengthen testing. The essential idea is first to use specification-based testing to discover all traversed program paths and then to use Hoare logic to prove their correctness. During the proof process, all errors on the paths can be detected.

TBFV is a specific specification-based testing approach that takes both the precondition and post-condition into account in test case generation [15]. To precisely describe this strategy, we first need to introduce functional scenario. $S_{pre}$ and $S_{post}$ denote the pre- and post-conditions of operation $S$. Let:

$$S_{post} = (G_1 \land D_1) \lor (G_2 \land D_2) \lor ... \lor (G_n \land D_n) \quad (1)$$

$G_i$ and $D_i$ ($i=1,..,n$) are two predicates, called guard condition and defining condition, respectively. The definition of functional scenarios and FSF (functional scenario form) are list below:

**Functional Scenario** = $S_{pre} \land G_i \land D_i$ \quad (2)

In the definition of functional scenario, $S_{pre} \land G_i \land D_i$ is treated as a scenario: when $S_{pre} \land G_i$ is satisfied by the initial state (or intuitively by the input variables), the final state (or the output variables) is defined by the defining condition $D_i$.

$$FSF = (S_{pre} \land G_1 \land D_1) \lor (S_{pre} \land G_2 \land D_2) \lor ... \lor (S_{pre} \land G_n \land D_n) \quad (3)$$

A systematic transformation procedure, algorithm, and software tool support for deriving an FSF from a pre-post style specification written in SOFL have been developed in our previous work [16]. First, generate test cases from specification. Second, form path triple and the definition are below:

$$\{S_{pre} \land G_i\} P \{Di\} \quad (4)$$

$P$ is called a program segment, which consists of decision (i.e., a predicate), an assignment, a return statement, or a printing statement.

Finally, repeatedly apply the axiom for assignment to derive a pre-assertion, denoted by $P_{pre}$. The correctness of the specific path is transformed into the implication $S_{pre} \land G_i \rightarrow P_{pre}$. If the implication can be proved, it means that no error exists on the path; otherwise, it indicates the existence of some error on the path.

IV. TBFV-M IN MBSE

Model-Based Systems Engineering (MBSE) combines process and analysis with architecture. In the development process using MBSE, as shown below, the users’ requirements are obtained first and the requirement document is usually written in natural language.

![Fig. 1. TBFV-M usage scenario](image)

To obtain requirements without ambiguities, we may generate a SysML Model. During the model-driven development process, we use the SysML Model Diagram to communicate with the user, because it does not contain many mathematical symbols and syntax.

During the Model-Driven process, model is an important medium for the Model based system engineering development.
The TBFV-M method is mainly used to verify whether SysML Activity Diagram model meets the user's requirements written in SOFL (Structured-Object-oriented-Formal Language).

V. PROCEDURE OF TBFV-M

The TBFV-M method takes the specification describing the users’ requirements and the SysML Activity Diagram model as input and verifies the correctness of the SysML model with respect to the specification. The procedure of TBFV-M is illustrated in Fig.2.

B. Functional Scenarios Derivation

The overall goal of functional scenario derivation is to extract all functional scenarios completely in "Spec ∧ Gi ∧ Di" form (FSF), as mentioned above in TBFV section. Because this part is not our main topic and has been researched before. In our work, we assume that an FSF of the specification has been available somehow. The below segment of the process buy ticket, mentioned previously, shows the FSF generated from the specification described in the last one.

```
1 buy_ticket_pre: age>0
G1: age<=6
D1: price = 0
2 buy_ticket_pre: age>0
G2: age>6
D2: price = 10
3 buy_ticket_pre: age<=6
```

C. Test Path Generation

A test path auto-generation tool based on the SysML Activity Diagram model takes the model as input and generates test cases as outputs automatically. First, we use transformation algorithm to compress the input Activity Diagram, which may contain unstructured module. The transformation is a cyclic process, dealing with loop module, concurrent module and the problem of multiple starting nodes separately. After compressing, we transform this unstructured activity diagram into an intermediate representation form Intermediate Black box Model (IBM). IBM consists of one basic module and a map from black box to the corresponding original actions. The third phase of our approach is test path generation based on IBM. Details of automated test paths generation algorithm and implementation of unstructured SysML Activity Diagram has been developed in our previous work [19].

The Loop module in the SysML activity diagram can be considered as a node collection, and these nodes in the collection can be cycled multiple times, as shown in Fig.3.

```
process buy_ticket {age: int; price: int}
pre: age>0
post: age<=6 and price = 0
or
age>6 and price = 10
end_process
```

Fig. 3. Classification of loop modules

The first step in the transformation algorithm of the Loop module is to identify the loop module, the second step is to compress it into a black box node loop, and finally reinsert it into the original SysML activity diagram. Fig.4 shows the process.
In the SysML activity diagram, the most common form of a concurrent module is a pair of fork node and join node and all actions between these two nodes, as shown in Fig. 5 (a). The logical representation is AND. However, the synchronization stream can also be the logical relationship OR, as shown in Fig. 5 (b). Depending on how many concurrent streams can be synchronized by the join node, the parallel modules can be divided into partJoin concurrent and noJoin concurrent, as shown in Fig. 5 (c) and (d) below, respectively.

On the test path generation algorithm for concurrent modules, the first step is to identify the concurrency module, the second step is to compress it into a black box node FJ (Fork-Join), and finally reinsert it into the original SysML activity diagram, as shown in the following Fig. 6.

For concurrent modules, we can use the Concurrent module path generation algorithm and generate the test path automatically. For the compressed basic path, the test path generation algorithm of the basic module can be applied. Once the basic path is generated, replace the FJ black box with the test path generated from the concurrency module.

After completing the initialization step, find a matching functional scenario for each element in edge list. The specific operation is: the edge after the integration compares with Spre Gi in the functional scenario, if exactly the same, then we find the edge with the matched functional scenario. If there is no exact matched functional scenario, then there is an inaccurate modeling problem and needs to be refined. Therefore, immediately terminate the program, the problem of the edge will also be returned. After traversing all the edge_list, we also need to check whether each in FS_list has been visited. If there is an unvisited functional scenario, then it means that there is a requirement that the model fails to be represented in the specification, and the model needs to be refined.

Establish Path Triple and apply each node with the axiom in Hoare Logic. “(S_{pre} ∧ G_1 ∧ D_1) (i = 2, ..., n)” denote one functional scenario and P = \{node_1; node_2; ...; node_n\} be a program path in which each node_i (i = 2, ..., n) is called a functional node, which is a DecisionNode, ActionNode, or others.

To verify the correctness of P with respect to the functional scenario, we need to construct Path Triple: \{S_{pre}\} P \{G_i ∧ D_i\}.
Each node has different processing approach, and the details are listed in the form below.

<table>
<thead>
<tr>
<th>Node Type</th>
<th>Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActionNode (assignment)</td>
<td>The axiom for assignment</td>
</tr>
<tr>
<td>ActionNode (input/output)</td>
<td>SKIP</td>
</tr>
<tr>
<td>Others node</td>
<td>SKIP</td>
</tr>
</tbody>
</table>

**TABLE I. PROCESSING APPROACH OF AD NODE**

F. Implication

Prove the implication. Finally, the correctness of one path whether it meets the corresponding requirement is changed into the proof of the implication \( S_{pre} \land G_i \rightarrow S_{pre} \). If the implication can be proved, it means that the path can model one part of the requirement; otherwise, it indicates the existence of some error on the path.

Formally proving the implication \( S_{pre} \land G_i \rightarrow S_{pre} \) may not be done automatically, even with the help of a theorem prover such as PVS, depending on the complexity of \( S_{pre} \) and \( P_{pre} \). Our strategy is as follows: if the complexity of data structure is not high, we will transform the problem into solver, which can achieve full automation. Otherwise, if achieving a full automation is regarded as the highest priority, as taken in our approach, the formal proof of this implication can be "replaced" by a test. That is, we first generate sample values for variables in \( S_{pre} \) and \( P_{pre} \), and then evaluate both of them to see whether \( P_{pre} \) is false when \( S_{pre} \) is true. If this is true, it tells that the path under examination contains an error.

VI. SUPPORTING TOOL

We have developed a prototype software tool to support the TBFV-M method. Specifically, it provides five major functions, which are functional scenario generation, test path generation, matching function scenarios to test paths, pre-condition derivation, verification of test paths, and output of verification result. The tool interface is shown in Fig. 7.

![Fig. 7. The process of transformation of concurrent modules](image)

VII. CASE STUDY

Now we show a motivation example to detail the process of MBSE and TBFV-M method described in the article above. First, we will get a requirement from the user, which consists of inform the description.

According to the specification, we can construct a set of SysML model and the Activity Diagram is shown below.

![Specication](image)

![Activity Diagram](image)

We can find the expression is described with SOLF. After getting ready with all the input, specification and Activity Diagram, we will start the TBFV-M method process. First, derive Functional Scenarios from specification and generate test paths from Activity Diagram. The result is shown as below.

![Test Path](image)

At the same time, we can extract data constraints from each test scenario, which is used for matching with functional scenario. Then, the matching process is shown below. If it does not exist a matched functional scenario, then it means that it exists a problem in the model, exactly in this unmatched test path. This path is not established accurately according to the requirements described in specification in the activity diagram model. If the match succeeds, it indicates that the test path is designed for the matched test scenario.
We will do the verification of test scenario according to the successfully matched functional scenario. First, we establish Path Triple and then apply the axiom of Hoare Logic to derive $P_{pre}$, pre-assertion of one path for the corresponding test path. The blow figure chose the forth path and matched the first functional scenario as an example and shows the substitution process, from bottom to up. So, the top one $\neg c \land b \land c = \neg c \land b \lor r = \neg b \lor r$ is the $P_{pre}$.

Finally, we turn this verification problem into proving whether the pre-condition of specification can imply $P_{pre}$. If it can be proved, means that the path satisfies the requirement. If not, there is a problem existing in the model, exactly in this unmatched test path. If the matched pre-condition can imply the corresponding $P_{pre}$ of all the test paths in the model, then the model is satisfied with the user’s requirements.

From the above segment, we can see the implication $(-c > 0 \land b \geq 0 \land \neg p = \text{FALSE} \land \neg c \leq \neg b \land c = \neg c \land b \lor r = \neg b \lor r)$ is true. This it means that the test path is satisfied with the corresponding functional scenario. We have proved all the test paths, due to the space limit, we omit further details.

VIII. EVALUATION

After finishing the supporting tool, we established 20 example cases to test our system. These test cases include 5 correct ones and the others include errors. All the incorrect Activity Diagrams fail to express the needs fully and correctly, such as missing some logic branch or having mistaken on some logic branch.

And the result is that the supporting tool has the ability to figure out these mistakes, as our expectation.

IX. CONCLUSION

We have presented an approach, known as TBFV-M (Testing-Based Formal Verification for Model), for requirement design error detection in SysML. Activity Diagrams by integrating test cases generation and Hoare Logic. The principle underlying TBFV-M is first to derive functional scenarios from specifications and generate test scenarios from Activity Diagrams. Then match them and verify each test scenario according to the corresponding functional scenario. Hoare logic is used during the verification process. TBFV-M method solve the limitation of TBFV, not concerning about models and solved the problem of inconsistent, incomplete, and inaccurate models. It has advantage in reducing the probability of system error and shortening the developing time.

X. ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number 26240008, and Defense Industrial Technology Development Program JCKY 2016212B004-2.

REFERENCES


