Automated Model-Based Test Case Generation for UML Activity Diagrams using EFSM

Nazanin Shahbazi
Data Mining Laboratory, Department of Computer Engineering
Alzahra University
Tehran, Iran
e-mail: N.shahbazi@student.alzahra.ac.ir

Mohammad-Reza Keyvanpour
Department of Computer Engineering, Faculty of Engineering
Alzahra University
Tehran, Iran
e-mail: keyvanpour@alzahra.ac.ir

Abstract— Model-based testing (MBT) is a promising approach for generating test cases from system models, providing high levels of automation and effectiveness. The potential for automation in MBT is possible if the model is formal and machine-readable. A commonly employed formal modeling technique is the representation of a system as an extended finite state machine (EFSM). However, in practice, formal models are not common in the industry. Activity diagrams, on the other hand, are well-suited for generating test cases, but their lack of formal semantics can lead to ambiguous interpretations and make them unsuitable for automation. In this paper, we propose an efficient approach that maps UML Activity Diagrams into EFSMs, providing a formal modeling of the system under test (SUT) and utilizes JUnit and ModelJUnit Java libraries to automatically generate test cases, using coverage measures to evaluate them. Our approach aims to overcome the limitations of automation in MBT, while achieving efficient coverage and execution time metrics.

Keywords—Model-Based Testing; Automatic Test Case generation; Activity Diagram; Extended Finite State Machine; ModelJUnit; JUnit

I. INTRODUCTION

Software engineering is a field that employs a methodical and structured approach to the development, operation, and maintenance of software systems [1]. The goal of software organizations is to develop and deliver software within specified timeframes and budgets [2]. However, software development, particularly for complex systems, is prone to human errors, necessitating the assessment of software for errors during source code generation [3]. Software testing plays a crucial role in identifying faults and is a vital but costly phase in software development and maintenance. One of the most challenging aspects of testing is generating test cases, which is essential to ensure the success of the testing phase [4].

Modeling techniques can enhance software testing productivity when used in this context [5]. Model-based test case generation (MB-TCG) is a methodology that involves the generation of tests using system models, which are based on model-based testing (MBT) [6]. Using MBT allows testing to begin early in the software development process, running tests without requiring access to the source code of the system under test, owing to its black-box nature. This technique uses system models to generate and execute test cases automatically, reducing testing time and effort [7].

Unified Modeling Language (UML) is a popular option for software modeling as it offers a high level of expressiveness [8]. Activity Diagram is the most used UML diagram in model-based test case generation [6]. UML Activity Diagram is a semi-formal specification that can be used to describe the workflow of the system and captures critical system behaviors [9]. Although this Diagram is widely used in this field, there is a lack of automated techniques for test generation from UML activity diagrams [10]. In addition to the mentioned problem, due to the lack of formal semantics, the use of UML diagrams can lead to inconsistency, transformation problems, and different interpretations [5]. Another issue of note is that UML specification presents a challenge in terms of navigation and comprehension [11].

Using formal models is one approach to mitigate these issues, as they offer precise semantics for representing system behavior [8]. Finite State Machines (FSMs) are a favorite notation in formal system modeling and testing software [5]. Although, formal models are rarely used in practice, probably because developers lack the necessary training and familiarity with the mathematical notation [11].

Formal models are highly recommended in Model-Based Testing (MBT) due to their ability to automate the testing process, resulting in improved efficiency and effectiveness [5]. Extended finite state machine (EFSM) is one of the formal models that have received significant attention and extensive study over the past few decades [12]. An EFSM model is an enhanced model based on FSM. This model can represent many complex systems containing both control and data parts [13].

In the context of this paper, we propose an efficient and systematic approach that maps UML Activity Diagrams into EFSM following some transformation rules. By using JUnit and ModelJUnit Java libraries, test cases are automatically generated from EFSM. The main contribution of this paper is the definition of transformation rules to map the various elements of the UML Activity Diagram into constructions of the Extended Finite State Machine (EFSM). Furthermore, to tackle the issue of complexity and understandability of the UML specification problem, we utilize an updated version of the metamodel for UML Activity Diagrams, specifically tailored for simplicity and applicability while ensuring a close alignment with the EFSM metamodel used in our research.

https://doi.org/10.51408/csit2023_05 26
The rest of this paper is organized as follows: Section II describes some related research. Afterwards, Section III presents UML Activity Diagram and Extended Finite State Machine. Section IV describes our approach and its implementation. In Section V, the results obtained by implementing the proposed approach on a sample scenario have been discussed. Finally, the conclusion and future work are placed in Section VI.

II. RELATED WORK

Numerous research studies have been conducted on the subject of generating test cases from UML Activity Diagrams. One of the most relevant related works in the area of automated test case generation from UML models is the article by Smith et al. [5]. They proposed an approach for generating test cases from UML sequence diagrams using the Extended Finite State Machines (EFSMs). Their approach showed promising results, but one of its weaknesses was that it only worked for sequence diagrams. We built our work upon the approach proposed by this article, but we extended it to work with UML activity diagrams (AD) since ADs are more commonly used in software development, as shown in a recent survey [6]. Our approach can generate test cases that cover a wider range of scenarios since Activity diagrams can model more complex behavior and allow for more varied interactions between objects than sequence diagrams [14].

In [15], the authors propose an approach for generating scenario-based test cases from UML Activity Diagrams. The proposed approach utilizes an intermediate model named Extended Activity Dependency Graph (EADG), which extends activity graphs to generate test scenarios. However, this approach differs from ours in that it does not utilize model-driven engineering (MDE) concepts or formal models for test generation.

In [16], the authors propose a novel method for generating test cases using UML Activity and Sequence Diagrams. Their approach involves the conversion of Sequence Diagram into a graph, which refers to Sequence Graph, and transforming the Activity Diagram into the Activity Graph. Test suite generation is achieved by merging the graphs into a single software graph. While this methodology shares similarities with ours, it diverges by not employing MDE concepts or formal models for test generation.

Other elements of Activity Diagrams defined by UML 2.5 are not in the scope of this article.

B. Extended Finite State Machine

The Extended Finite State Machine (EFSM) is a well-known formal specification technique that is commonly employed to define the various states and actions of a software system. This method is widely used to describe the behavior of software systems in a precise and unambiguous manner [18].

An EFSM can be formally represented by a 6-tuple \((s_0, S, V, I, O, T)\) where \(S\) is a finite set of states with initial state \(s_0\); \(V\) is a finite set of context variables; \(I\) is a set of transition inputs; \(O\) is a set of transition outputs; and \(T\) is a finite set of transitions.

Each transition \(t_x \in T\) can also be represented formally by a tuple \(t_x = (s_i, s_j, p_{tx}, A_{tx}, i_{tx}, o_{tx})\), where \(s_i, s_j\) are the origin and target states of transition \(tx\), and \(i_{tx}, o_{tx} \in I\) represents the input parameters of the beginning of the transition \(tx\), such as events that can be interpreted as special types of input parameters, and \(o_{tx} \in O\) denotes the output results at the end of the transition \(tx\). \(P_{tx}\) represents the predicate conditions (guards) with their respective context variables, and \(A_{tx}\) denotes the operations (actions) with their respective current variables. EFSM models can be represented as a directed graph \(G(V, E)\). The elements of \(V\) represent the states of an EFSM, and \(E\) denotes its transitions [11].
IV. PROPOSED APPROACH

This section presents an automated approach for generating test cases from UML activity diagrams using EFSM. The methodology involves several steps that begin with translating the UML activity diagram into a formal EFSM model using Atlas Transformation Language. The generated EFSM model is used to automate the test generation process, where test cases are generated based on EFSM-based methods. These methods provide coverage for different paths and states within the system. ModelJUnit and JUnit libraries are used to facilitate the generation of executable test cases. Acceleo is then used to perform a Model-To-Text (M2T) transformation, converting the generated test cases from a model representation into a textual representation that can be executed.

A. Metamodeling Constructs

In order to establish a clear and structured foundation for our research, we defined two metamodels: the UML Activity Diagram metamodel, which serves as a source for our model transformation process, and the Extended Finite State Machine metamodel, which represents the target model. To implement these metamodels, we utilized the Eclipse Modeling Framework (EMF) and represented them in Ecore.

The official UML specification [8] can be complex to navigate and hard to understand. Therefore, this metamodel has been heavily criticized, and the use of simplified metamodels is prevalent in most of the literature on this subject [11]. To address this issue, a simplified metamodel for the Activity Diagram is proposed in this study, depicted in Fig. 1, in comparison to the metamodel specified by the OMG.

The proposed metamodel contains 19 metaclasses and eliminates constructs that are not frequently used in practice. By streamlining the metamodel, it is easier to understand and apply in practice, providing a more practical and straightforward approach to Activity Diagrams.

As shown in Fig. 2, the metamodel used for EFSM is comprised of six metaclasses, with EFSM serving as an abstraction of an Extended Finite State Machine. Within the EFSM entity, there are states, transitions, and context variables [5].

B. Metamodels Transformation Principles

A detailed description of the rules for transforming an Activity Diagram into an Extended Finite State Machine is provided in this section. The defined transformation rules are listed below:

- RInitialNode: For the node of type InitialNode, first an EFSM is created with the same name of the AD, then initial state S0 is added to it. The initial state is used to update both the previous state and the current state.

- RDecisionNode: RDecisionNode rule operates in two stages, depending on the type of decision node being processed. For each conditional decision node, this rule creates one new state and two new transitions. The transition with a guard of "true" leads to the newly created state, while the transition with a guard of "false" returns to the previous state. For switch decision nodes, this rule creates a new state and transition for each guard. Regardless of the type of decision node, each transition has output, guard, and action labeled with the guard of the specific control flow, and any output pin from the node before the decision node is treated as the transition's input. After creating each new state, this rule updates both the previous and current states.

- RMergeNode: This rule handles merge nodes in the Activity Diagram (AD) by creating a new transition for each node that has a control flow leading to the merge node. These transitions do not have any output, guard, or action, but any output pin of the nodes connected to the merge node is considered to be a corresponding transition's input. The transitions all lead to the state that was already created for the
node after the merge node using the "RActivityNode" rule. After creating these transitions, both the previous and current states are updated.

- **RSignalActivityNode**: For nodes of type ExecutableActivityNode, SendSignalActionNode, and ReceiveSignalActionNode, this rule creates a new state and transition. The new state is connected to the previous state using the transition, and all pins of the corresponding node in the AD are considered as inputs for the transition. Additionally, this rule checks all control flows of a node to see if any flow goes backward without any conditions. If such flows exist, a new transition is created to connect the corresponding states to each other, and the inputs of these transitions are labeled as previously explained. After creating these new states and transitions, the previous and current states are updated as usual.

Our implementation of the transformation rules relied on Atlas Transformation Language (ATL).

In this study, we are utilizing a slightly modified version of some of the lazy rules proposed in [5], in combination with our matched rules. Specifically, the following lazy rules are employed:

- **LrInitialState**: The initial state $S_0$ is created, followed by incrementing the state order. Both the previous state and the current state are updated to the newly created initial state. Additionally, the name of the Activity Diagram is stored in a variable.

- **LrState**: A new state is generated, the state order is incremented, the previous state is updated to the current state, and the current state is changed to the newly created state.

- **LrTransition**: A transition is established, connecting the previous state to the current state. The transition's input is labeled with the input/output pin of the Activity Diagram. The output, guard, and action can be null and depend on the type of the activity node.

In our study, we used an automatic four-step approach for generating test cases. This approach involves implementing the EFSM model interface, implementing the adapter, generating the test cases, and then concretizing them. These four steps were automatically generated using Acceleo.

V. EXPERIMENT

This section presents an overview of our proposed approach, as well as the results and a discussion of the experiment.

A. Case Study

This section describes a case study demonstrating the practical application of our approach. The UML Activity Diagram in Fig. 3, depicts an ATM (Automatic Teller Machine) withdrawal process. Initially, we created an Activity Diagram model that described the behavior of the system, using the Activity Diagram editor implemented in the EMF. To transform this Activity Diagram model into an executable model, we utilized the transformation rules implemented in ATL.

After the execution of the transformation rules, we obtained an EFSM model that reflected the behavior of the system. Fig. 4 illustrates the resulting EFSM model that was generated. We created this model using the Visual Paradigm Drawing Tool by utilizing the XML output file generated from executing the transformation rules on the ATM Activity Diagram.

![ATM Activity Diagram](image)

![ATM EFSM Model](image)

Fig. 3. ATM Activity Diagram

Fig. 4. ATM EFSM Model

Afterwards, we followed the 4-step automatic approach, as previously described, to generate test cases using classes AtmModel, AtmAdapter, AtmTest, and AtmJUnit. These classes were obtained through the implementation of the four steps in the approach. Finally, our approach was put to the test through a series of test cases designed to assess its effectiveness. To execute the tests, we utilized the Eclipse Modeling Framework (EMF) in conjunction with the ATMTest and ATMJUnit classes.
Reflecting the importance of coverage criteria to test engineers when creating the test cases [19], we used three distinct coverage types to evaluate the application of our approach. A test model can be represented by an abstract regular set \( P \subseteq \Sigma^a \), where \( \Sigma \) is an alphabet denoting possible actions, and \( a \in N \cup \{\ast\} \) indicating that each test can have any length. Considering the fact that \( P \) may be huge or even infinite, including \( 'a' \) in \( N \cup \{\ast\} \) means that tests can vary in length within specific boundaries [20]. As a result, coverage criteria can be outlined as in Definition 1.

**Definition 1 (coverage).** A set of tests \( S \subseteq P \) is said to cover a test-model \( P \subseteq \Sigma^a \) under the coverage criteria \( C = \{C(i)\}_{i \in I} \), \( C(i) \subseteq \Sigma^a \) if \( \forall i \in I : (C(i) \cap P \neq \emptyset \Rightarrow C(i) \cap S \neq \emptyset) \) and \( \bigcup_{i \in I} C(i) = \Sigma^a \), where the index set \( I \) is used to name the coverage sets and the coverage sets \( C(i) \) are used to identify the aspects that are covered by the tests [20].

During testing, we tracked state, action, and transition coverage metrics for each individual test case, as presented in Table I. The state coverage metric indicates the number of states visited, and its formula can be expressed as (1), where \( |SV| \) is the number of states visited during testing and \( |TS| \) is the number of total states.

\[
\text{state coverage} = \left( \frac{|SV|}{|TS|} \right) \times 100
\]  

The action coverage metric reflects the number of actions performed during testing. The action coverage formula is given by (2), where \( |AP| \) is the number of actions performed during testing and \( |TA| \) is the number of total actions.

\[
\text{action coverage} = \left( \frac{|AP|}{|TA|} \right) \times 100
\]  

The transition coverage metric measures the number of transitions that were visited. The transition coverage formula is shown in (3), where \( |TV| \) is the number of states triggered during testing and \( |TT| \) is the number of total transitions.

\[
\text{transition coverage} = \left( \frac{|TV|}{|TT|} \right) \times 100
\]

Prior to initiating the tests, we established initial values for the card attribute (111), psw attribute (123), and balance attribute (100.00) belonging to the Bank class of the SUT.

**B. Obtained Results**

The generated test cases and the results obtained from their execution are all reported in Table I. We have also included a visually informative bar chart, shown in Fig. 5 to illustrate the metrics associated with each test case.

Fig. 5 displays the coverage metrics for each test case generated using our approach. It is important to note that the possible number of states, actions, and transitions that can be visited may vary depending on the specific test case and its associated parameters.

In particular, for the initial set of generated test cases, the total number of states in the ATM system is determined to be 14 based on the EFSM representation in Fig. 2. However, since the value is not generated within the scope of the test cases T1, T2, and T3, the actual number of states that can be visited is reduced. According to our EFSM analysis, only 5 states are possible to be visited, given the absence of the value.

**C. Discussion**

Our approach’s novelty lies in its ability to automate the entire testing process, with the exception of manually defined stubs. To the best of our knowledge, the level of automation we have achieved with ADs has not been reported before, making our approach a valuable contribution to the field. To further illustrate the superior automation achieved by our approach, we conducted a comprehensive comparison with three of the most automated and recent works in the field, as shown in Table II.

**TABLE I. GENERATED TEST CASES**

<table>
<thead>
<tr>
<th>Test case</th>
<th>card</th>
<th>psw</th>
<th>value</th>
<th>State</th>
<th>Action</th>
<th>Trans.</th>
<th>Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>111</td>
<td>123</td>
<td>5</td>
<td>14/12</td>
<td>60/144</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>111</td>
<td>246</td>
<td>5</td>
<td>14/12</td>
<td>60/144</td>
<td>13.6</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>222</td>
<td>246</td>
<td>2</td>
<td>14/12</td>
<td>24/144</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>111</td>
<td>123</td>
<td>50</td>
<td>14/12</td>
<td>144/144</td>
<td>16.7</td>
<td></td>
</tr>
<tr>
<td>T5</td>
<td>111</td>
<td>123</td>
<td>200</td>
<td>14/12</td>
<td>144/144</td>
<td>23.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>14/12</td>
<td>144/144</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE II. COMPARISON OF THE LEVEL OF AUTOMATION**

<table>
<thead>
<tr>
<th>Approach</th>
<th>Level of Automation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Hettab et al. [15]</td>
<td>Moderate</td>
<td>Test data can only be generated manually from the automatically generated test scenarios.</td>
</tr>
<tr>
<td>Meiliana, I. et al [16]</td>
<td>Fairly High</td>
<td>Testing itself is automated, but the approach necessitates the manual creation of 2 UML diagrams.</td>
</tr>
<tr>
<td>Our approach</td>
<td>High</td>
<td>Automates the entire testing process except for stubs.</td>
</tr>
</tbody>
</table>

Fig. 5. The state, action, and transition coverage, along with the execution time (in milliseconds), for each generated test case.
VI. CONCLUSION AND FUTURE WORK

In this paper, we proposed an automated model-based test case generation approach using Activity Diagrams. To formalize the activity diagrams, we proposed a simplified metamodel that was used to map ADs to an EFSM representation using some transformation rules. We then applied a four-step approach to generate test cases, using JUnit and ModelJUnit to execute the tests. Our approach was able to achieve good coverage and time results, demonstrating the effectiveness of our methodology.

As future work, other UML diagrams, such as state and class diagrams, can be incorporated into the proposed test generation process. Moreover, advanced testing techniques, such as mutation testing, can be used to enhance the generated test suite. Finally, our approach can be applied to a real-world software system to further evaluate its effectiveness and scalability.

REFERENCES


