Test Cases Generation for Object-oriented Software from Decision Slicing of UML Activity Diagram

Wasiur Rhmann1* and Vipin Saxena1

1Department of Computer Science, B. B. Ambedkar University (A Central University), Lucknow, U.P., India.

Authors’ contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JSRR/2016/26872

Editor(s):
(1) Vito Di Maio, Institute of Cybernetics "E. Caianiello" CNR, C / O Complex Olivetti, Via Campi Flegrei, 34, Italy.
(2) Luigi Rodino, Professor of Mathematical Analysis, Dipartimento di Matematica, Università di Torino, Italy.

Reviewers:
(1) Boni Garcia, Universidad Rey Juan Carlos, Spain.
(2) Yulia Kartika Gunadi, Petra Christian University, Indonesia.
(3) Mukesh Mann, Gautam Buddha University, India.

Complete Peer review History: http://www.sciencedomain.org/review-history/15675

ABSTRACT

Software testing is an integral part of the software development life cycle. Design of good test cases is a key challenge in software testing. Test cases can be designed from different artifacts like requirements, design and software code. In Software engineering, different UML diagrams are used for designing and analysis of the software systems. The main contribution of this work is to propose a novel technique of test cases generation from UML activity diagram using an iterative method. Iterative methods are used in numerical analysis for generation of solution of equation iteratively. In the present work a test cases generation technique from decision slicing of UML activity diagram is presented. Decision slices for each decision nodes are derived from the Activity Flow Graph (AFG) of the activity diagram. Test cases are generated for each activity path of the activity flow graph. Decision nodes at each activity path are used to generate system of equations and these equations are solved by an iterative method to generate test data for each activity path. A case study of ticket purchasing from ticket vending machine using UML activity diagram is presented.
Keywords: Software testing; program slicing; activity diagram; iterative method; test cases.

1. INTRODUCTION

Software testing is a creative and challenging part of the software development life cycle. Quality and reliability of the developed software products is ensured with the help of software testing. Software testing cost generally depends on the how efficiently testing is performed. There are mainly three activities in test process: test cases generation, test case execution and test evaluation. Generation of test cases is most critical activity. A test case is the triplet $[I, S, O]$ where $I$ is input data, $S$ is the state of the system and $O$ is expected output [1-2].

There are mainly two approaches of software testing namely: white box and black box. In white box testing, there is need to know internal details of programs and in black box testing test cases are generated with the requirement specification documents and internal structure of program is not considered. In black box testing, test cases are designed with consideration of the functionality of the software system. Model based testing is a kind of black box testing which is used to find bugs earlier in the software development life cycle. In model based testing expected models which capture requirements of the software under test are created and a test tool is used to automatically derive the test cases from the designed models [3].

Large systems are more complex to test. To reduce the development and testing costs of modern software, software engineers usually advocate the use of object-oriented analysis and design paradigm [4]. Testing object oriented system with the code is tedious and complex task. Object Management Group (OMG) [5] introduced the standardized Modeling Language called Unified Modeling Language [6] to facilitate to use different diagrams to present static and dynamic views of the systems. Activity diagram focuses on the flow of behavior of a system [10]. Different UML diagrams like class diagram, use case diagram and activity diagram are used to represent system at different level of abstraction. Activity diagram is like flow chart but it is more expressive than flow chart. Sequential or concurrent control flow can be represented by an activity diagram. UML activity diagram is used by several researchers as initial specification of the system for generation of test cases [11-13].

Large and complex software can be decomposed into manageable pieces of code to easily handle the complex software and there are accurate dependency exists among different pieces of the software. Program slicing techniques are used to handle the complexity of programs which arises due to large size of programs. Program slices extracts are those statements of the programs which are relevant to a particular computation [14]. Dynamic slices are often smaller size than static slices as it includes only those statements of program which are executed for particular input data.

Program slicing techniques are used in different aspects of the software development like debugging [15], software maintenance [16], software measurement [17] and software testing [18]. In the present work, authors used a new slice method of UML activity diagram called decision slice and generated test cases for each activity path. Test data corresponding each activity path is generated using an iterative method of numerical analysis. Iterative method is used in [19] for generating tests data in the context of program code. In our approach iterative method is used at design level for test cases generation. The rest of the paper is organized as follows: Section 2 of the paper describes the UML activity diagram and iterative method used for test cases generation. Section 3 presents the methodology used for test cases generation. A case study of automatic ticket
In the current work, UML activity diagram is used to represent the functionality of the system. Activity diagram contains mainly two types of nodes namely action nodes and control nodes. Action nodes contains Activity, Action, SendSignal and AcceptEvent while Control nodes contains InitialNode, FinalNode, Decision, Merge, Fork and Join. There are several levels of activity modeling: Basic, Intermediate, Complete and Structured Complete and extra structure activities in UML 2.1 superstructure [20]. Test scenarios are used to check whether all business process related to the software are tested end to end. Test cases and test scenarios are often used by Software tester synonymously [21]. An activity diagram (AD) as defined in [22] is based on seven tuple described below:

\[ AD = (A, T, F, J, R, a_0, a_f) \]

where

- \( A \) is finite set of activities \( a_0, a_1, \ldots, a_n \), representing nodes,
- \( T \) is finite set of transitions \( t_0, t_1, \ldots, t_n \), representing (edges),
- \( F \) is finite set of forks \( f_0, f_1, \ldots, f_n \),
- \( J \) is finite set of join \( j_0, j_1, \ldots, j_n \),
- \( R \subseteq (A \times T) \cup (T \times F) \)
- \( A \) is the flow relation, \( a_0 \) is initial activity node and \( a_f \) is final activity node.

3. PROPOSED METHODOLOGY

The steps in the proposed methodology are described below:

1. Draw UML activity diagram;
2. Convert the activity diagram into Activity Flow Graph (AFG);
3. From AFG generate activity paths;

**AFG for UML activity diagram is drawn, for each component of activity diagram there is a node in AFG. Decision nodes are represented by circle and for fork/join rectangle are used.**

Activity paths are generated by traversing the activity flow graph in DFS manner and if there is fork node then BFS are used in traversing the nodes of activity flow graph.
4. For each activity path, test data is generated. The decision nodes of each activity paths are used to form the equations;

4.1 Decision slice Decision slice of a decision node (DN) on path P is the set of nodes of AFG, upon which the outcome of decision node depends directly or indirectly.

4.2 Then compute the input dependency set for each decision node on activity path. A subset of input (I) variables along an activity path (P) on which decision node (DN) outcome depends is called input dependency set ID (DN, I, P).

4.3 Formulate the decision functions from decision nodes. Decision functions are formulated based on input variables on which that decision node depends. Suppose if a decision node i depends on input variables x, y and z.

\[ F_i: a_i x + b_i y + c_i z + e_i \]

Then decision function will be \( F_i \); \( a_i x + b_i y + c_i z + e_i \) with respect to their variables evaluate the decision function at the current input \( I'_k = (i_1, \ldots, i_j, \ldots, i_m) \) and at \( I'_k + (0, \ldots, \Delta i_j, \ldots, 0) \) and compute the divided difference by using the formula given below

\[
F(I'_k + (0, \ldots, \Delta i_j, \ldots, 0)) - F(I'_k) / \Delta i_j
\]

This gives the coefficient of \( i_j \) in the linear function for the decision function \( F \) corresponding to node \( i \) in \( P \). Similarly other coefficients are computed.

Decision functions are also formulated from values of each decision nodes on activity path; If a decision nodes have no input variables then decision function for that decision node will contain one Boolean variable with 0 for false and 1 for true value of the function.

4.4 Formulate the linear equations from decision nodes.

Convert linear functions into inequalities. For conversion of linear functions into inequalities we choose >= and < operators. If a decision node on an activity path needs to be true for traversal of that path then relational operator will be same as used in decision node else it will be reverse of the decisional nodes.

Compute decision residuals. Decision residual of a decision node for an input is the value of decision function computed by executing its decision slice at that input;

Then apply the relaxation technique and formulate the inequalities of increment of the input variables and convert these into equations and solve these equations. On solving these equations we get the increment values which will give the next input values. This process of iteration continues till we get the values which traverse the activity path.

4. A CASE STUDY OF TICKET PURCHASING FROM AUTOMATIC TICKET VENDING MACHINE

In the present section UML activity diagram is designed for ticket purchasing from automatic ticket vending machine and test cases are generated using presented approach.

Fig. 1 presents UML activity diagram for purchasing tickets from ticket vending machine. There are three partitions which are Passenger, Ticket Vending Machine and Bank. Passenger can perform different activities. Passenger selects the ticket types and destination of travel which inputs the price of ticket. Then passenger selects number of ticket types (n). Then ticket vending machine computes the total amount and displays the total fare. Then machine process for payment. Ticket vending machine accepts money either by card or cash. If passenger pays with card then bank is used for authorization of payment and ticket is issued. If passenger inserts cash then machine calculates the difference between inserted money and total fare. If money inserted is less than the total fare then a message of insufficient amount is displayed. If difference between total fare and money inserted is 0 then ticket is issued and money inserted is more than the total fare then machine issues ticket along with the returned money.

Fig. 2 represents the activity flow graph of activity diagram given in Fig. 1. In Fig. 1 there are 21 different constructs like start, end node, activities, fork, join and decision node these are converted into nodes of activity flow graph. Nodes 8, 13 and 15 represent decision nodes of activity diagram. Nodes 17 and 20 represent the fork and join nodes of activity diagram respectively.

Activity paths from activity flow graph are generated by traversal the graph in depth first
search manner and there is a fork and join node
then node between them are traversed in
breadth first search manner which is activity
path4. Generated activity paths are given below:

Activity Path1 1-2-3-4-5-6-7-8-9-16-21
Activity Path2 1-2-3-4-5-6-7-8-9-10-12-13-
15-16-21
Activity Path3 1-2-3-4-5-6-7-8-9-10-12-13-
14-21
Activity Path4 1-2-3-4-5-6-7-8-9-10-12-13-
15-17-18-19-20-21

Below the steps of test case generation for
activity path2 are given:

4.1 Derivation of Decision Functions

Let x for the decision nodes which do not
contains constraints and values of x are 0 or 1.

Activity path 2 contains three decision nodes 8,
13 and 15. For each decision nodes presented in
the corresponding activity path, we identify the
input variables which will affect the outcome of
the decision node. It is given by

\[ ID(\text{DN}_1, P) = x \]  \hspace{1cm} (5)
\[ ID(\text{DN}_2, P) = (p, n, a) \]  \hspace{1cm} (6)
\[ ID(\text{DN}_3, P) = (p, n, a) \]  \hspace{1cm} (7)

For each decision nodes available in the
the corresponding activity path authors formulated
the linear equations. Activity path2 has three
decision nodes so there will be three equations:

\[ F_1 = a_1 x + e_1 \]  \hspace{1cm} (8)
\[ F_2 = a_2 p + b_2 n + c_2 a + e_2 \]  \hspace{1cm} (9)
\[ F_3 = a_3 p + b_3 n + c_3 a + e_3 \]  \hspace{1cm} (10)

Fig. 1. UML activity diagram for ticket purchasing from automatic ticket vending machine
Here equation 8 represents decision node with no input variables so it takes the form of equation 11:

$$F_1=x$$  \hspace{1cm} (11)

We select $x=1$ for traversal of activity path 2.

Decision functions corresponding equations 9 and 10 are formulated from values of corresponding decision nodes on activity path and are given by equations:

$$F_2=b-0$$  \hspace{1cm} (12)

$$F_3=b-0$$  \hspace{1cm} (13)

Let $I_0=(p, n, a)=(100, 2, 220)$ then input will be $(x, p, n, a)=(1, 100, 2, 220)$, this will not traverse the activity path 2.

Then next input is obtained using the above iterative method:

We take increment $\Delta p=2$ and computed the values of $F_1$ at $I_0$.

and at $(p_0, n_0, a_0)+((\Delta p, 0, 0))=(100, 2, 220)+(2, 0, 0)=(102, 2, 220)$.

Then divided difference $F_1(p_0+\Delta p, n_0, a_0) - F_1(p_0, n_0, a_0)/\Delta p$ is calculated which gives the value of $a_1$

$$a_1=(16-20)/2=-2$$

Similarly other values are calculated and recorded in Table 1.

### Table 1. Values of coefficient of equations

<table>
<thead>
<tr>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$c_1$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>-2</td>
<td>-100</td>
<td>-100</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Values of $F_2$ and $F_3$ given in equations 12 and 13 are computed at $I_0$ and recorded in Table 2.

### Table 2. Calculation of functions with the input values

<table>
<thead>
<tr>
<th>$a$</th>
<th>$p$</th>
<th>$n$</th>
<th>$F=b=a-c=a-p*n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td>100</td>
<td>2</td>
<td>$F_2=20$</td>
</tr>
<tr>
<td>220</td>
<td>100</td>
<td>2</td>
<td>$F_3=20$</td>
</tr>
</tbody>
</table>

Constant terms $e_i$ in the equations 9 and 10 are computed by executing the decision slices at $I_0$. Then the values of input $I_0$ and coefficient of equations recorded in Table 1 are put in the linear equations 9, 10 and equated with the corresponding values computed in Table 2.

$$a_1100+2b_1+220c_1+e_1=20$$

$$-200-200+220+e_1=20$$

$$e_1=200$$

Similarly value of $e_2$ is calculated and the values of $a_i$, $b_i$, $c_i$ and $e_i$ and corresponding functions 9 and 10 are recorded in the Table 3.

### 4.2 Computation of Decision Residuals

Decision slices corresponding each decision nodes on the activity path are executed with
current program inputs \( I_0 \) and values of decision functions are evaluated.

**Table 3. Derivation of decision functions**

| \( a_1 \) | -2 | \( a_2 \) | -2 |
| \( b_1 \) | -100 | \( b_2 \) | -100 |
| \( c_1 \) | 1 | \( c_2 \) | 1 |
| \( e_1 \) | 200 | \( e_2 \) | 200 |
| \( F_2 \) | \(-2p-100n+a+200\) | \( F_3 \) | \(2p-100n+a+200\) |

Decision residual at \( I_0 \) for decision nodes on Activity Path2 are recorded in Table 4:

**Table 4. Computation of decision residual**

| \( I_0 \) | Residual |
| (100, 2, 220) | \( R \left( DN_{i_0}, AP_{2}\right) = 20 \) |
| (100, 2, 220) | \( R \left( DN_{i_0}, AP_{2}\right) = 20 \) |

Construction of a system of linear constraints and to solve them to obtain increments for the current input:

Linear arithmetic representation of decision nodes are converted into inequalities:

Decision nodes corresponding decisional functions \( F_2 \) and \( F_3 \) which are given in Table 2 needs to be true for traversal of Activity path2 So we have taken the same relational operator as used in these decisional nodes and inequalities corresponding each nodes is written in equations 15, 16.

\[
-2p-100n+a+200 \geq 0 \quad (15)
\]

\[
-2p-100n+a+200 = 0 \quad (16)
\]

Then relaxation technique is applied to equations 15 and 16 and decision residual computed in Table 3. Set of constraints on increments of \( p \), \( n \) and \( a \) are given in equations 17 and 18.

\[
-2\Delta p-100\Delta n+\Delta a \geq -20 \quad (17)
\]

\[
-2\Delta p-100\Delta n+\Delta a = -20 \quad (18)
\]

**4.3 Converting the Inequalities into Equations**

Inequalities are converted into equations by introducing the variable \( u \):

\[
-2\Delta p-100\Delta n+\Delta a-u = -20 \quad (19)
\]

\[
-2\Delta p-100\Delta n+\Delta a = -20 \quad (20)
\]
Table 5. Test cases number, test paths and input data corresponding each test path

<table>
<thead>
<tr>
<th>Test case no.</th>
<th>Test path</th>
<th>Input data (x, p, n,a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC1</td>
<td>Select ticket Type-&gt;Select Destination(p) -&gt; Select Number of Tickets(n) -&gt; c=p*n -&gt; Display Amount to be inserted -&gt; Process Payment-&gt; Authorize Card Payment -&gt; Eject Ticket -&gt;end</td>
<td>(0,0,0,0)</td>
</tr>
<tr>
<td>TC3</td>
<td>Select ticket Type-&gt;Select Destination(p) -&gt; Select Number of Tickets(n) -&gt; c=p*n -&gt; Display Amount to be inserted -&gt; Process Payment-&gt; Insert Money(a) -&gt; Calculate -&gt; b=a-c -&gt; Insufficient Amount-&gt;end</td>
<td>(1, 100, 3, 200)</td>
</tr>
<tr>
<td>TC2</td>
<td>Select ticket Type-&gt;Select Destination(p) -&gt; Select Number of Tickets(n) -&gt; c=p*n -&gt; Display Amount to be inserted -&gt; Process Payment-&gt; Insert Money(a) -&gt; Calculate -&gt; b=a-c -&gt; Eject Ticket -&gt;end</td>
<td>(1, 100, 2, 200)</td>
</tr>
<tr>
<td>TC4</td>
<td>Select ticket Type-&gt;Select Destination(p) -&gt; Select Number of Tickets(n) -&gt; c=p*n -&gt; Display Amount to be inserted -&gt; Process Payment-&gt; Insert Money(a) -&gt; Calculate -&gt; b=a-c -&gt; Eject Ticket -&gt; Return Money -&gt; end</td>
<td>(1, 100, 2, 2010)</td>
</tr>
</tbody>
</table>

On solving the equations 19 and 20, we found the values of \(u\), \(\Delta a\), \(\Delta p\), \(\Delta n\) given below:

\[
\begin{align*}
\Delta a &= -20, \quad \Delta p = 0, \quad \Delta n = 0 \\
\end{align*}
\]

New values of \(x\), \(p\), \(n\), \(a\) are obtained by adding the values of \(\Delta a=-20, \Delta p =0, \Delta n=0\) to \((x, p, n, a)\).

\((x, p, n, a) = (1, 100, 2, 200)\).

Here computed new values lead to traversal of the desired activity path and computed new approximation of test inputs which traverse an activity diagram is obtained from previous approximation of the solution and it's residual. This technique will be used iteratively to obtain new input values until desired activity path is traversed.

Similarly test data for other test paths are calculated and recorded in Table 5.

Java programming language is used for implementation of the proposed approach. For the given problem there are two decision nodes 13 and 15 which take input values. Java program contains functions of the decision nodes 13and 15 of the programs. And Initial input is taken in the program as \((p, n, a) = (100, 2, 220)\).

Then decision node symbol is provided based on the path traversal then our program generates test data corresponding that path. Fig. 3 shows the test data for test path 2.

5. COMPARISON WITH THE RELATED WORKS

Swain et al. [24] presented a technique of test case generation using UML 2.0 sequence diagram. Authors constructed a Message Dependency Graph from sequence diagram. Conditional predicates are selected from message dependency graph by traversing the graph. Slices are computed for each conditional predicates and test cases are generated for each slice. Test data is generated by satisfying all constraints corresponding each slice. Swain et al. [25] used condition slicing and generated test cases from UML interaction diagram. Authors used message guards of interaction diagram and created conditioned slice for each message guards. For each conditioned slices test cases are generated. This approach is advantageous when number of messages in sequence diagram is in large number. Li et al. [26] used extenics theory for generation of test cases from UML activity diagram. Authors converted the UML activity diagram into Euler circuit and applied the Euler circuit traversal algorithm to generate test cases. Generated test cases consist of test cases sequences there is no input data.

Jena et al. [27] used UML activity diagram for test paths generation and test cases are generated from Activity flow graph of the activity diagram by traversing the graph in Depth First Search manner. Technique presented by authors only generates test paths not exactly test cases. Kundu and Samanta [28] presented an approach of test cases generation from UML 2.0 with use
case scope. Authors considered test coverage criteria called activity path coverage criteria. Test cases generated from their approach can detect faults like synchronization faults, faults in loops. Their technique will not be efficient when the activity diagram will be larger.

Samuel and Mall [29] used dynamic slicing technique on the flow graph of activity diagram then sliced the diagram for each conditional predicates and test cases generated for each slice. In their technique if there are large number of conditional predicates then there will be a lot of test cases corresponding each conditional predicated while in our technique conditional predicates on each path are considered simultaneously to generate test data which will reduce the number of test cases. In their approach authors do not check each valid test path. Their technique generates two test cases for each condition while our technique generates only one test case for each condition or decision nodes. In Fig. 4 number of generated test cases from our technique and technique presented by Samuel are compared. In the presented case study there are four test cases generated while by using Samuel technique there will be 6 test cases corresponding each conditional node in the flow graph as there are 3 conditional nodes in flow graph.

![Number of Test Cases](image)

**Fig. 4. Number of test cases by our technique and technique presented by Samuel**

% reduction in number of test cases will be=$\frac{2}{6} \times 100 = 33.33\%$

6. CONCLUSIONS

Slicing was initially developed for software programs. It helps to manage the complexity of the large programs. In the present work, test cases are generated from UML activity diagram using an iteration method. For each test path generated from AFG of the UML activity diagram there is test data which executes that test path. Our approach used a decision slicing criteria which slices the activity diagram based on decision nodes of the activity diagram. Then iterative method is used to generate the test data corresponding each path of the activity diagram. This method is an innovative method which uses the iterative method of numerical analysis for generation of test cases from UML activity diagram. Most of the techniques used by researchers generate only test paths of activity diagram while our technique generates test paths along with the input data. Our approach generates the reduced number of test cases and test data generated by our approach satisfies the activity paths. In future research other UML diagrams may be used with the iterative methods for test data generation. Presented relaxation based technique which is used to solve the equations of test paths may be useful in identifying the infeasible test paths from UML diagrams and identification of infeasible paths from UML may cut down the testing cost drastically.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES


