The Effect of Program and Model Structure on MC/DC Test Adequacy Coverage

ICSE '08: Proceedings of the 30th international conference on Software engineering

Ajitha Rajan  
Mats P.E. Heimdahl  
Dept. of Comp. Sci. and Eng.  
University of Minnesota

Michael W. Whalen  
Advanced Technology Center  
Rockwell Collins Inc.
MC/DC as a coverage metric for testing

- MC/DC widely used in critical systems such as in avionics or military
- Paper states MC/DC criteria can be “cheated” and heavily depends on code structure
What is MC/DC?

Modified Condition/Decision Coverage
=> Source code metric for measuring the quality of a test suite

Condition Coverage + Decision Coverage + Each condition must independently affect the outcome of a decision

Test Suite

Implementation

```c
int myFunc (bool c1, bool c2, bool c3)
{
    bool d1 = c1 or c2;
    bool d2 = d1 and c3;
    if (d2)
        return 1;
    else
        return -1;
}
```
MC/DC example

```c
int myFunc (bool c1, bool c2, bool c3) {
    bool d1 = c1 or c2;
    bool d2 = d1 and c3;
    if (d2)
        return 1;
    else
        return -1;
}
```

Test Suite
Subset of all possible input tuples which satisfies MC/DC criteria

For Example: \{TFF, FTF, FFT, TTT\} // (c1 c2 c3)
Problems with MC/DC

Same program written in a different way (*d1* has been **inline**d)

```c
int myFunc (bool c1, bool c2, bool c3)
{
    bool d2 = (c1 or c2) and c3;
    if (d)
        return 1;
    else
        return -1;
}
```

Previous test suite does not satisfy MC/DC criteria anymore!

If correct expression should have been

```c
bool d2 = (c1 and c2) and c3;
```

bug will not be revealed by current test suite!

**Problem:** A test suite satisfying MC/DC criteria would have detected fault, which shows that MC/DC coverage can be affected by program structure.
Case examples

**Goal:** show that test suite providing MC/DC over non-inlined version will achieve lower MC/DC over implementation that is inlined.

Case examples used in industry
- Aircraft Display Window Manager (3)
- Flight Guidance System (3)

Toy examples
- Wheel Brake System
- Sensor Voting Example

Systems were available as Simulink Models
Experiment Setup

1. Translation framework used to generate different implementations
2. Test suite generated through NuSMV model checker
3. Obtain “minimal” test suite using a (naive) algorithm
4. Compare measured/achievable MC/DC

- **Achievable MC/DC**: Complete coverage sometimes not possible (e.g., masking)
- **Measured MC/DC**: Coverage provided by test suite
### Size of Generated Test Suites

How many tests are needed for each implementation to achieve MC/DC?

<table>
<thead>
<tr>
<th>Function</th>
<th>Non-Inlined</th>
<th>Inlined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Reduced</td>
</tr>
<tr>
<td>DWM_1</td>
<td>180</td>
<td>18</td>
</tr>
<tr>
<td>DWM_2</td>
<td>299</td>
<td>39</td>
</tr>
<tr>
<td>DWM_3</td>
<td>2522</td>
<td>23</td>
</tr>
<tr>
<td>ToyFGS_05</td>
<td>4445</td>
<td>75</td>
</tr>
<tr>
<td>Latctl</td>
<td>315</td>
<td>52</td>
</tr>
<tr>
<td>Vertexmax</td>
<td>1415</td>
<td>235</td>
</tr>
<tr>
<td>WBS</td>
<td>271</td>
<td>10</td>
</tr>
<tr>
<td>Sensors</td>
<td>103</td>
<td>10</td>
</tr>
</tbody>
</table>

**DWM_3 requires 20 times more tests to cover inlined version!**

=> Mostly Boolean logic, leading to complex expressions in the inlined implementation
Achieved Coverage (MC/DC)

- Generated test suite for non-inlined version of DWM_3 achieves very low MC/DC on inlined implementation (13%)
- Interesting: ToyFGS_05 has a much lower achievable MC/DC in inlined version => many DNF expressions containing redundancy causing strong masking effect
- Inadequacy ranging from 13% to 86%, statistically supported on a 5% significance level (including industrial examples only)
Conclusions

MC/DC is indeed highly sensitive to structure of implementation!

• **Suggestions**
  - Different coverage metric that takes masking into account (independent of code structure)
  - Apply coverage on model domain instead of code domain

• **Problems with experiments**
  - Small number of examples
  - Test suite reduction too naive

• **Personally**
  - Removing toy examples from statistics is questionable
  - Effectiveness of MC/DC?