Evaluating Testing Methods by Delivered Reliability

Frankl, Hamlet, Littlewood, Strigini
IEEE TOSE Aug98

Theta

- $\Theta$ - failure probability for a randomly drawn point
- $\mathbb{E}(\Theta)$ – expected value of failure probability

3.3 single failure region, with sub

- Debug with subdomains
  - $\mathbb{E}(\Theta) = q\Pi(1-d)^T$
- Operational
  - $\mathbb{E}(\Theta) = q(1-q)^T$

Triangle Code

```
cin >> a >> b >> c;
type = "scalene"
if (a == b || a == c || b == c) type = "isosceles";
if (a == b && b == c) type = "equilateral";
if (a+b+c|b>=a+c|c>=a+b) type = "not a triangle";
if (a<=0|b<=0|c<=0) type = "bad inputs";
cout<< type
```

Regression Testing

- When is it done?
- Why is prioritization important?
Rate of fault detection

- How is this different than Theta (Θ)?

Granularity

- Fine
- Coarser

Mutation Coverage

- Insert “standard” faults to create set of mutants
- A mutant is “killed” if its output is different than “oracle”
- Run test cases until all mutants are killed
- Or, calculate how many mutants are killed by each test case

Hypothesis

- What was their expectation?

Table 1 - techniques

<table>
<thead>
<tr>
<th>Technique</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Figure 1
APFD

\[ \text{APFD} = 1 - \frac{TF_1 + TF_2 + \ldots + TF_n}{nm} + \frac{1}{2n}. \]

Figure 2

Figure 4