

誤りシミュレーションによるテスト評価手法の提案*

4Q-4

原 圭吾 小山 徳章†

株式会社東芝 研究開発センター S & S 研究所†

1 はじめに

近年、ソフトウェアは大規模かつ複雑になってきている。それに伴ってソフトウェアのテストも大規模化しており、テストケースの作成や実行にかかるコストは非常に大きなものとなっている。そのため、誤り (bug) 検出効果の高いテストケースを抽出して優先的に投入したり、テスト結果を自動的に解析することができれば、大きくコストを削減することができると思われる。

本稿では、テスト/デバッグ作業の効率化を目的として、テストケースの誤り検出能力に関する情報を生成し、その情報によって誤りの存在する可能性についての評価を行う手法を提案する。本手法は2つのフェーズからなる。まずテストケース分析フェーズでは、SDL[1] 等で記述された仕様に人為的に誤りを埋め込んで仮想的に実行を行うことで、誤り検出のための情報を生成する。続いてテスト結果分析フェーズでは、前フェーズで得られた情報と実際のテスト結果を比較集計し、実際に誤りが存在する可能性についての評価を行う。

2 テストケース分析フェーズ

テストケース分析フェーズでは、誤りの埋め込まれた仕様 (誤り仕様と呼ぶ) を仮想的に実行させ、誤り検出能力に関する情報を生成する。分析手続きの概要を以下に示す。

1. テスト対象の仕様から、想定している誤りごとに誤り仕様を生成する。想定している誤りの情報は予めデータベースに用意しておく。図2は、図1の仕様に対する誤り仕様の例である。
2. 与えられたテストケースと仕様 (元々の仕様とそれぞれの誤り仕様) の組み合わせ全てについて、仕様言語インタプリタ等を用いて仮想的な実行を行う。
3. 元々の仕様の実行結果と、各誤り仕様の実行結

果とを比較する。比較結果が異なる誤り仕様とテストケースの組み合わせについては、実行結果を表に記録する。この表を誤り情報テーブルと呼ぶ。表1は誤り情報テーブルの例である。表1において tc1~tc5 はテストケースを、m1~m4 は誤り仕様に埋め込まれた誤りを表す。

誤り情報テーブルの欄が空であるかどうかは、次のような意味を持つ。

- 欄 tc,m が空: 元々の仕様と m 誤り仕様について、テストケース tc を適用した場合の実行結果が同じであり、 tc では誤りによる差異が検出できないことを表す。
- 欄 tc,m が *result*: 元々の仕様と m 誤り仕様について、テストケース tc を適用した場合の実行結果が異なっており、かつ m 誤り仕様の実行結果が *result* であることを表す。これは、 tc が m 誤りによる差異を検出するのに有効であることを示す。

例えば表1の m4,tc2 に対応する欄は、テストケース tc2 について仕様 S と m4 誤り仕様の実行結果が異なっており、m4 誤り仕様の結果が「テストケース tc2 中の命令2で不合格と判定された」ことを表す。さらにこれは、誤り m4 の検出に tc2 が有効であることを表す。また表1より、テストケース tc1 が誤り m1~m4 の検出にあまり有効でないことや、誤り m2 を検出するのに有効なテストケースが無いこと等を知ることができる。

表1: 誤り情報テーブル

	tc1	tc2	tc3	tc4	tc5
m1				命令5 不合格	
m2					
m3			命令2 不合格	命令2 不合格	命令4 不合格
m4		命令2 不合格			命令2 不合格

*Evaluation of Software Test Results with Enbugged Simulation

†Keigo HARA, Noriaki KOYAMA

‡Systems & Software Research Laboratories, Research & Development Center, Toshiba Corporation

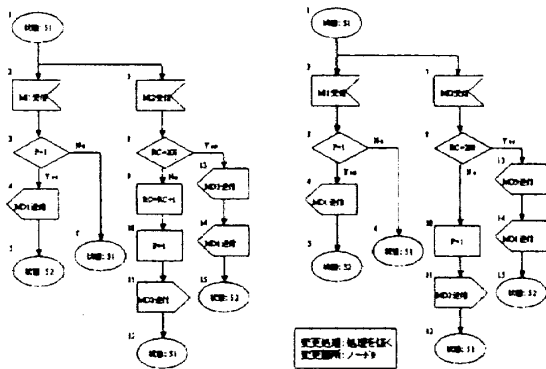


図 1: ソフトウェア仕様 S 図 2: 誤り仕様の例

3 テスト結果分析フェーズ

テスト結果分析フェーズでは、前フェーズによって得られた情報から、誤りが実際に存在する可能性についての評価を行う。手続きの概要を以下に示す。

1. 誤り情報テーブルを参照し、それぞれ誤りの検出に有効なテストデータを誤り評価データに記録する。例えば表 1 の誤り情報テーブルの場合、tc2 は誤り m4 に対して有効であるので、誤り評価データの m4、「有効」の欄に tc2 を記録する。
 2. 与えられたテストケース集合とテスト対象ソフトウェアを用いてテストを行い、テスト結果を記録する。表 2 は、テスト結果の一例である。
 3. テストケースの一つに着目し、対応するテスト結果と誤り情報テーブルの列を比較する。比較した結果、誤り情報テーブル中に一致する欄があれば、そのテストケースを誤り評価データの対応する欄に記録する。これを全てのテストケースについて行う。例えば表 1 と表 2 において tc4 に着目した場合、テストケース tc4 のテスト結果は、誤り情報テーブルの m1,tc4 の欄と一致する。このとき、誤り評価データの行 m1 の「結果一致」の欄に tc4 を記録する。
 4. 誤り評価データの各誤りについて、誤りの存在する可能性についての評価を記録する。
 - 有効なテストケースの集合が空である場合、評価結果は - とする。
 - 上記以外で、有効なテストケースの集合と実行結果の一致したテストケースの集合が一致した場合、評価結果は A とする。
 - 上記以外で、結果の一致したテストケースの集合が空でない場合、評価結果は B とする。
 - 上記以外で、結果の一致したテストケースの集合が空の場合、評価結果は C とする。
- 誤り評価データの例を表 3 に示す。

誤り評価データには、それぞれの誤りが実際に存在する可能性についての評価が記録される。各評価の意味を以下に示す。

- 評価結果が - の場合、有効なテストケースがなく、その誤りについて評価が出来ないことを表す。
- 評価結果が A の場合、その誤りが実際に存在する可能性が高いことを表す。
- 評価結果が B の場合、その誤りが存在する可能性はあるものの、A ほどではないことを表す。
- 評価結果が C の場合、その誤りが存在しない可能性が高いことを表す。

表 3 より、m1~m4 の各誤りの存在する可能性は m1 が最も高く、逆に m3 は存在しない可能性が高いことがわかる。これらの情報により、誤りを特定する際に評価の高い誤りから検討していくことができ、作業効率を上げることができると考えられる。また、m2 の誤りについては有効なテストケースがなく、比較的テストが手薄な部分と言える。このように、手薄な部分を明らかにすることで、手薄な部分を補う作業についても有効であると考えられる。

表 2: テスト結果の例

tc1	tc2	tc3	tc4	tc5
合格	合格	合格	命令 5 不合格	命令 2 不合格

表 3: 誤り評価データの例

	有効	結果一致	評価結果
m1	tc4	tc4	A
m2			-
m3	tc3, tc4, tc5		C
m4	tc2, tc5	tc5	B

4 おわりに

本稿では、ソフトウェアのテストにおけるテストケースおよびテスト実行結果を評価し、誤りの特定のための情報を付加することで、テスト/デバッグのコストを軽減する手法を提案した。本手法は、ソフトウェアが与えられた仕様を満たしているかどうかをテストする場合に有効であり、発生しやすい誤りや致命的となりうる誤りが予め想定されている場合に特に有効である。

今後は、本手法を実現するためのシステムを試作しての実験評価を行う。また、誤り評価データの評価手法についても実験評価の結果をもとに改善検討を続けていきたい。

参考文献

[1] 若原恭, 長谷川晴朗. 仕様記述言語 SDL. 株式会社カットシステム, 1996.