

Automated Static Analysis of Unit Test Code

Rudolf Ramler, Michael Moser and Josef Pichler

2016 IEEE 25 ~ 28page

背景

- テストコードは製品のコードとともに進化
- テストコードにも障害は含まれる
- 自動化された単体テストは広く使用されているソフトウェア品質保証手段

背景

- テストコードは現在軽く扱われている
- ソフトウェアシステムの継続的な開発、メンテナンスフェーズを通してテストコードも絶えず適合、および維持する必要がある

目的

- 単体テストコードの品質を分析および改善するための支援として、**自動静的分析**のアプリケーションについての報告

分析方法

- 分析の対象はJUnit（テストフレームワーク）
- 静的分析を実施
- よく発生する問題を 4 種類の問題領域に区別
 - (a) xUnitテストフレームワークの使用
 - (b) ユニットテストの実装
 - (c) テストスイートのメンテナンス
 - (d) SUTのテストの容易性

(a) xUnitテストフレームワークの使用

- テストフレームワークの変化に伴い，新たなライブラリやモックフレームワークを導入することが必要
- インターフェースの複雑さや要領を得ない説明文によって混乱や問題が発生

(b) ユニットテストの実装

- そもそもユニットテストの有効性は実際の動作と予想される動作の逸脱を検出するので、**誤ったテスト設計**，論理テストを**誤ったテスト実装**に転送
- 戻り値のアサートの欠落が最も頻繁に発生する問題

(b) ユニットテストの実装

- チェックが多すぎると、実行環境にわずかな変更があることによってテストが失敗したり、異なる順序で実行すると失敗したりする

(c) テストスイート

- テスト中のシステムの変更が「迅速な修正」の場合、テストコードをメンテナンスする作業は余分は無意味な負担とみなされる
- 自動テストは上記のようにメンテナンスがなされないことが原因で一貫性がなく理解しにくい状態になり、最終的に放棄される

(d)SUTのテストの容易性に関連する問題

- 外部依存関係の存在が挙げられる
- 実際の動作はテストで制御できないサービスの影響を受ける

静的分析

- 静的分析はコーディング規約と標準の遵守をチェックするプロセスを自動化
- PMDを利用したが42の追加ルールを実装
- よくテストされているJFreeChart5が対象
- 対象の中のorg.jfree.chart.junitパッケージの28個のテストクラスに実施

PMDの追加ルール

- PMDには300を超えるルールが備わっている
- 今回の目的の問題領域に関するものが少ない
- それぞれの問題に合わせた追加ルールを計42個追加

PMDの追加ルールの内訳

- アサートステイトメントに関するもの(15個)
- 命名規則に関するもの(8個)
- テストケースのロジックに関連するもの(6個)
- セットアップティアドアウンルーチンの実装に関連するもの(8個)
- その他(5)

実験結果・考察(1)

- 追加した**42**のルールによって検出された違反が合計で**982**個
- 検出された違反はアサートステイトメントに関連するものが**57%**（**559**個）を占める
- アサーションが正しく行われていないと確認された

実験結果・考察(2)

- アサートがメッセージを提供するかどうかをチェックするルールがあり，分析されたテストでは提供していなかった
- このルールで**280**の違反を示した
- 優先度の低い違反であるといえるので「推奨」と位置付ける必要がある

実験結果・考察(3)

- 命名規則では**292**の違反が見つかった
- 標準とわずかに異なる規則を使用していた（テストクラスには最後に s をつける等）
- これによって違反とはいえないものまで違反となっていたと考えられる

実験結果・考察(4)

- 残りの違反の多くは実際の問題を示している
- 追加したルールで検出したい問題とは異なる問題に関連する違反を発見
- 今回の実験とは関係のない違反が結果に含まれている

まとめ

- 単体テストを自動で行うコードは軽んじられていて、品質基準やメンテナンスが行われていない
- 単体テストの価値と利点が損なわれている
- 違反の三分の一が実際の障害

改善点

- 実践的な経験に基づくレビューが単体テストスイートを確認するのに最も重要な手段だと考えられる
- レビューは時間のかかる手動タスクであるので静的分析の向上も手段の一つ
- ルールセットの微調整の実施

所感

- 最初に述べられた4つの問題領域と追加したルールセットがどれに対応しているか書いておらず、問題領域を提示した意味が薄れていると感じた
- 今後の課題が多く発見されたという点が評価できると考えられる