

ミューテーションスコアを利用した テストセット評価システム

斉藤孝志 大久保弘崇 粕谷英人 山本晋一郎
(愛知県立大学大学院 情報科学研究科)

発表の構成

- ? 背景
- ? 目的
- ? ミューテーション解析
- ? テストセット評価システム
- ? まとめ

背景

ソフトウェア開発において、プログラムの
テストは不可欠
完全にテストできているのか?
テストセットは信用できるものか?

どのようなテストが有効か？

カバレッジ分析
全ての行、分岐、経路などが
実行されてるかを確認



テスト漏れを防止

カバレッジの数値が
基準値以上のテストセット ⇒ 安心??

目的

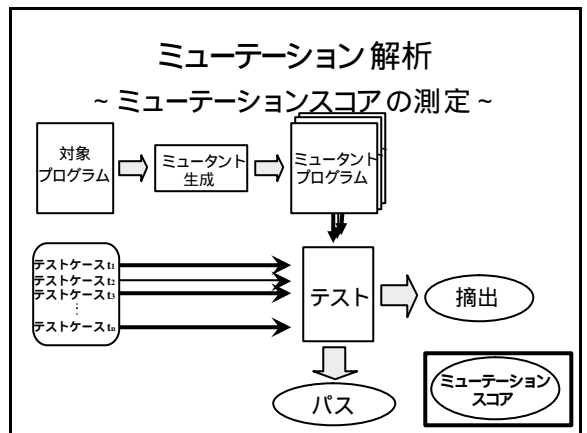
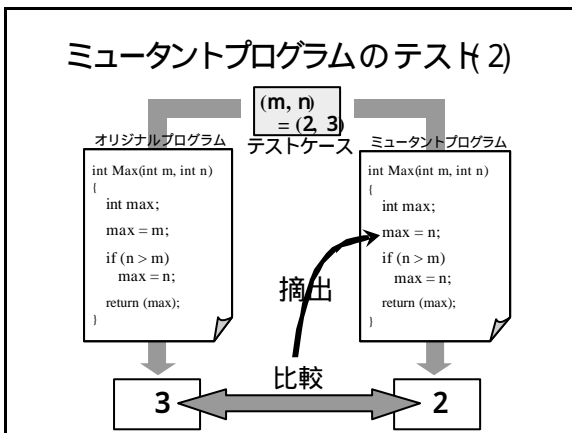
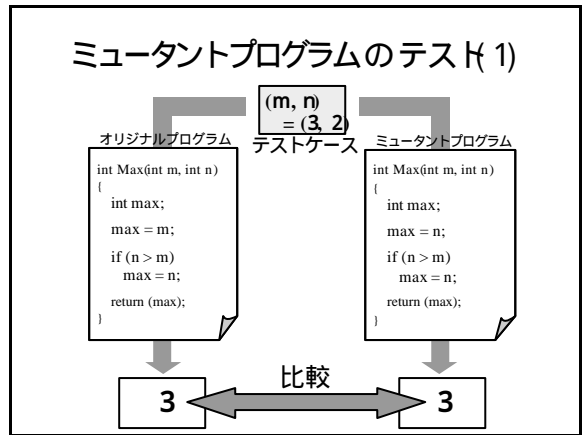
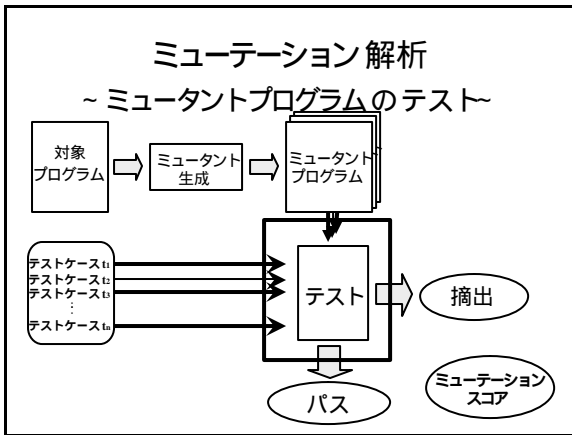
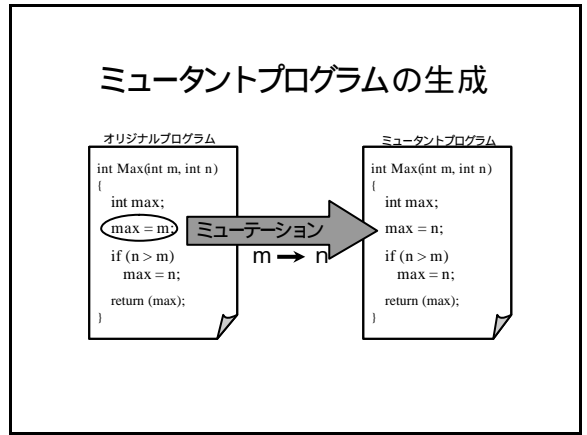
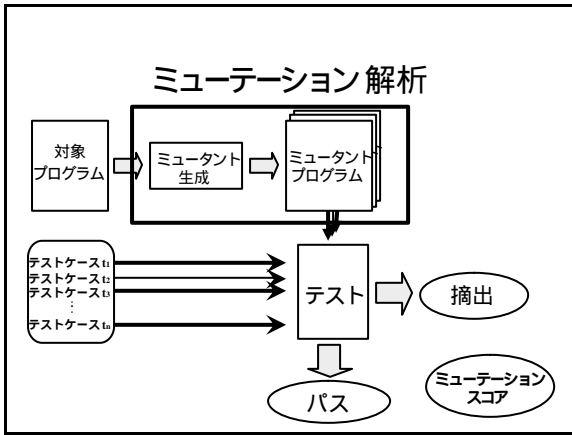
テストセットの品質を評価

意図的に混入したバグを含む
プログラムをテスト
抽出できるか

ミューテーション解析

ミューテーション解析

1993年 Untch, Offutt, Harrold によって提案

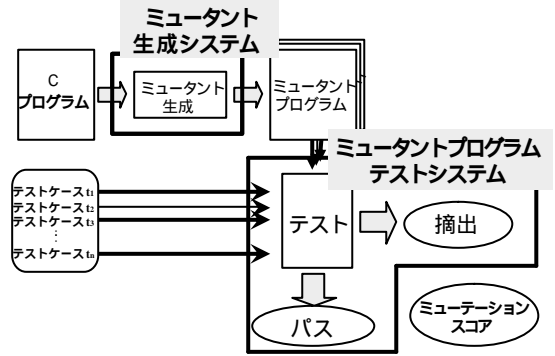


ミューテーションスコア

テストセットの品質を表す指標

$$\frac{\text{抽出したミュータントの数}}{\text{生成したミュータントの総数}} \times 100 [\%]$$

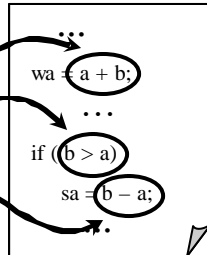
テストセット評価システム



ミュータント生成システムの対象ステートメント

対象ステートメント

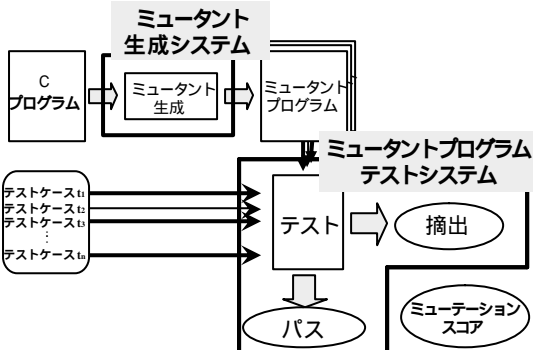
- 代入式の右辺
- if文の条件式



ミューテーションタイプ

- ? 変数
スコープ内にある同じ型の変数への置き換え
- ? 定数
定数: C C + 1, C - 1, 0
- ? 算術演算子 (+, -, *, /, %)
他の算術演算子への置き換え
- ? 関係演算子 (>, >=, <, <=, ==, !=)
他の関係演算子への置き換え

テストセット評価システム

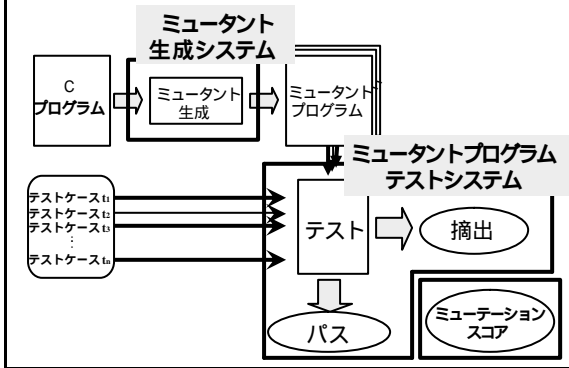


ミュータントプログラムテストシステム

1. ミュータントプログラムの生成 **抽出**
2. ミュータントプログラムのテスト
 - オリジナルプログラムのテスト結果と異なる
 - 正常時のテストの2倍以上の時間がかかる

強制的に終了

テストセット評価システム



ミューテーションスコアの測定

gzip-1.2.4 ? 8,163行 304KB ? テストケース: 2個	gawk-3.0.1 ? 23,638行 896KB ? テストケース: 40個
--	--

gzip-1.2.4のテストセットの ミューテーションスコア

- ? 生成したミュータントの総数 ... 4,001個
(10sec)
- ? 対象ステートメント ... 618個
- ? 抽出したミュータントの数 ... 774個
(55min)
- ? 強制終了させたもの ... 18個

$$MS = \frac{74}{401} = 19.3\%$$

gawk-3.0.1のテストセットの ミューテーションスコア

- ? 生成したミュータントの総数 ... 11,122個
(sec)
- ? 対象ステートメント ... 1,727個
- ? 抽出したミュータントの数 ... 個
()
- ? 強制終了させたもの ...

$$MS = \frac{??}{1122} = ??\%$$

まとめ

- テストセット評価システムの構築
- 既存のテストセットのミューテーションスコアの測定
gawk-3.0.1: ???%, gzip-1.2.4: 19.3%

今後の課題

- ミューテーションタイプの拡張
プログラマが犯しやすいミュータントの生成
- 他のオープンソースのテストケースの
ミューテーションスコアの測定