

When Do Changes Induce Fixes?

Jacek Sliwerski
Thomas Zimmermann
Andreas Zeller

60206050 大東誠弥

目次

- 背景
- 修正箇所 の 識別方法, 分析手法
- リンク数 と トランザクション
- 修正箇所 の 平均サイズ
- バグ と 各曜日 の 関係性
- まとめ

背景

- システムの状態を知るためにはプログラムの規則性を見る必要がある
- バグ情報にはプログラムの変更点と時間が記録されている

修正箇所推定の手順

- データベースよりバグを確認
- プログラム内の変化の確認
- バグが報告される前の,過去の変更点の確認

修正箇所の識別方法

- 変更点が修正されたものか確認する



開発者の間ではコメントにバグと報告番号を記載している

- トランザクションとバグの間のリンクを確認する

構文の分析手法

- バグデータのリンクの発見のため、ログやコメントを**証拠**と照らし合わせる



- バグ番号
- コード上の簡単な数値
- キーワード

リンク内容の分析

- リンクの状態により,リンク内の数値が変化



- バグが少なくとも1度は修正されている
- バグレポートの内容がログ,コメントと一致
- バグと修正を行った人物が一致
- トランザクションによって影響を受ける箇所
がバグに含まれている

リンク数とトランザクション

syn / sem	0	1	2	3	4	total
0	270 (1%)	1,287 (5%)	2,057 (8%)	1,439 (6%)	2 (0%)	5,055 (20%)
1	324 (1%)	4,152 (16%)	9,265 (37%)	1,581 (6%)	5 (0%)	15,327 (61%)
2	110 (0%)	1,922 (8%)	2,421 (10%)	482 (2%)	0 (0%)	4,935 (19%)
total	704 (3%)	7,361 (29%)	13,743 (54%)	3,502 (14%)	7 (0%)	25,317 (100%)

Table 1: Distribution of links accross different classes of syntactic and semantic confidence levels in ECLIPSE

syn / sem	0	1	2	3	4	total
0	560 (1%)	2,899 (5%)	4,281 (8%)	639 (1%)	8 (0%)	8,387 (16%)
1	1,211 (2%)	9,059 (17%)	16,336 (30%)	2,241 (4%)	22 (0%)	28,669 (54%)
2	478 (1%)	5,250 (10%)	9,133 (17%)	1,645 (3%)	12 (0%)	16,518 (31%)
total	2,249 (4%)	17,208 (32%)	29,750 (55%)	4,525 (8%)	42 (0%)	53,574 (100%)

Table 2: Distribution of links accross different classes of syntactic and semantic confidence levels in MOZILLA

修正箇所の平均サイズ

	fix-inducing	$\neg$ fix-inducing	all
fix	3.82 ± 26.32	2.08 ± 7.42	2.73 ± 7.87
$\neg$ fix	11.30 ± 63.02	2.77 ± 14.94	3.81 ± 26.32
all	7.49 ± 44.37	2.61 ± 13.66	3.52 ± 22.81

Table 3: Average sizes of fix and fix-inducing transactions for ECLIPSE

	fix-inducing	$\neg$ fix-inducing	all
fix	5.79 ± 37.37	2.12 ± 9.74	4.39 ± 30.05
$\neg$ fix	4.61 ± 30.59	1.91 ± 10.30	3.05 ± 21.39
all	5.19 ± 34.12	1.97 ± 10.13	3.58 ± 25.23

Table 4: Average sizes of fix and fix-inducing transactions for MOZILLA

修正箇所の平均サイズ

	fix-inducing	$\neg$ fix-inducing	all
fix	3.82 ± 26.32	2.08 ± 7.42	2.73 ± 7.87
$\neg$ fix	11.30 ± 63.02	2.77 ± 14.94	3.81 ± 26.32
all	7.49 ± 44.37	2.61 ± 13.66	3.52 ± 22.81

Table 3: Average sizes of fix and fix-inducing transactions for ECLIPSE

	fix-inducing	$\neg$ fix-inducing	all
fix	5.79 ± 37.37	2.12 ± 9.74	4.39 ± 30.05
$\neg$ fix	4.61 ± 30.59	1.91 ± 10.30	3.05 ± 21.39
all	5.19 ± 34.12	1.97 ± 10.13	3.58 ± 25.23

Table 4: Average sizes of fix and fix-inducing transactions for MOZILLA

修正箇所の平均サイズ

	fix-inducing	$\neg$ fix-inducing	all
fix	3.82 ± 26.32	2.08 ± 7.42	2.73 ± 7.87
$\neg$ fix	11.30 ± 63.02	2.77 ± 14.94	3.81 ± 26.32
all	7.49 ± 44.37	2.61 ± 13.66	3.52 ± 22.81

Table 3: Average sizes of fix and fix-inducing transactions for ECLIPSE

	fix-inducing	$\neg$ fix-inducing	all
fix	5.79 ± 37.37	2.12 ± 9.74	4.39 ± 30.05
$\neg$ fix	4.61 ± 30.59	1.91 ± 10.30	3.05 ± 21.39
all	5.19 ± 34.12	1.97 ± 10.13	3.58 ± 25.23

Table 4: Average sizes of fix and fix-inducing transactions for MOZILLA

バグと各曜日の関係性

% of revisions	Day of Week							avg
	Mon	Tue	Wed	Thu	Fri	Sat	Sun	
$P(\text{fix})$	18.4	20.9	20.0	22.3	24.0	14.7	16.9	20.8
$P(\text{bug})$	11.3	10.4	11.1	12.1	12.2	11.7	11.6	11.4
$P(\text{bug} \cap \text{fix})$	4.6	4.8	4.6	5.2	5.6	4.5	4.5	4.9
$P(\neg \text{bug} \cap \neg \text{fix})$	74.9	73.5	73.5	70.8	63.4	78.1	76.0	72.7
$P(\text{bug} \text{fix})$	25.1	22.9	23.3	23.5	23.2	30.3	26.4	23.7
$P(\text{bug} \neg \text{fix})$	8.2	7.1	8.1	8.8	8.7	8.4	8.6	8.1

Table 5: Distribution of fixes and fix-inducing changes across day of week in ECLIPSE

% of revisions	Day of Week							avg
	Mon	Tue	Wed	Thu	Fri	Sat	Sun	
$P(\text{fix})$	42.5	46.5	49.7	45.9	48.4	50.2	61.1	48.5
$P(\text{bug})$	39.1	44.1	41.2	40.8	46.2	44.9	26.4	41.5
$P(\text{bug} \cap \text{fix})$	19.4	23.6	22.8	21.6	26.9	19.6	13.2	21.9
$P(\neg \text{bug} \cap \neg \text{fix})$	37.8	33.0	31.9	34.9	32.3	24.5	25.7	31.9
$P(\text{bug} \text{fix})$	45.7	50.8	45.8	47.1	55.6	39.1	21.6	45.2
$P(\text{bug} \neg \text{fix})$	34.1	38.3	36.7	35.5	37.3	50.6	33.9	38.1

Table 6: Distribution of fixes and fix-inducing changes across day of week in MOZILLA

バグと各曜日の関係性

% of revisions	Day of Week							avg
	Mon	Tue	Wed	Thu	Fri	Sat	Sun	
$P(\text{fix})$	18.4	20.9	20.0	22.3	24.0	14.7	16.9	20.8
$P(\text{bug})$	11.3	10.4	11.1	12.1	12.2	11.7	11.6	11.4
$P(\text{bug} \cap \text{fix})$	4.6	4.8	4.6	5.2	5.6	4.5	4.5	4.9
$P(\neg \text{bug} \cap \neg \text{fix})$	74.9	73.5	73.5	70.8	63.4	78.1	76.0	72.7
$P(\text{bug} \text{fix})$	25.1	22.9	23.3	23.5	23.2	30.3	26.4	23.7
$P(\text{bug} \neg \text{fix})$	8.2	7.1	8.1	8.8	8.7	8.4	8.6	8.1

Table 5: Distribution of fixes and fix-inducing changes across day of week in ECLIPSE

% of revisions	Day of Week							avg
	Mon	Tue	Wed	Thu	Fri	Sat	Sun	
$P(\text{fix})$	42.5	46.5	49.7	45.9	48.4	50.2	61.1	48.5
$P(\text{bug})$	39.1	44.1	41.2	40.8	46.2	44.9	26.4	41.5
$P(\text{bug} \cap \text{fix})$	19.4	23.6	22.8	21.6	26.9	19.6	13.2	21.9
$P(\neg \text{bug} \cap \neg \text{fix})$	37.8	33.0	31.9	34.9	32.3	24.5	25.7	31.9
$P(\text{bug} \text{fix})$	45.7	50.8	45.8	47.1	55.6	39.1	21.6	45.2
$P(\text{bug} \neg \text{fix})$	34.1	38.3	36.7	35.5	37.3	50.6	33.9	38.1

Table 6: Distribution of fixes and fix-inducing changes across day of week in MOZILLA

まとめ

- トランザクションとバグのリンクを見ることで、修正箇所がわかる
- 修正を引き起こすトランザクションはプログラム中に多く存在する
- 曜日によってバグや修正の数が変わる