

不確かさに着目した revert コミット分析

村岡 北斗^{1,a)} 亀井 靖高^{1,b)} 佐藤 亮介^{1,c)} 鷗林 尚靖^{1,d)}

概要: ソフトウェア工学において「不確かさ」をマネジメントした開発を行うことは重要である。そのためには、実際の不確かさについて明らかにする必要がある。本論文では、Git で管理しているリポジトリを対象に revert されたコミットに含まれる不確かさがどのようなものであるか、目視での分析を行った。revert コミットと revert によって内容が打ち消された reverted コミットの 2 種類のコミットについて不確かさを調査した。1) 20 の OSS リポジトリ 570,086 コミットの中に不確かさを含む revert コミットは 23 件、不確かさを含む reverted コミットは 68 件存在した。2) 不確かさを含む revert コミット、reverted コミットともに、レベル 1 の認知的な不確かさが多い傾向にあった。それぞれの特徴としては、不確かさを含む revert コミットには、ソフトウェアの構造に不確かさが含まれてるコミットが多く、不確かさを含む reverted コミットの revert 理由は誤った解決策を取り消したのことが多い傾向にあった。

キーワード: 不確かさ, ソフトウェア開発, Git, revert

1. はじめに

ソフトウェア開発中の「不確かさ」と呼ばれる問題は、開発工程の様々な所に存在する。不確かさの例として、はっきりとしない要求、曖昧な仕様や、不明な API 仕様、修正内容が正しいかわからないなどといった問題が挙げられる。不確かさは、どのように扱っていいか一般的な対応方法がないため、開発者にとって扱いにくいもので一時的な措置をとられやすい。その結果として、不確かさはバグやソースコードの煩雑化の原因となりがちである。開発者は、不確かさを上手く管理したソフトウェア開発を行うことが必要とされている [2]。

近年、不確かさを抱擁したソフトウェア開発は重要な研究課題として、ソフトウェア工学において注目されている項目の 1 つである。ゴールモデリング、UML モデリング、モデル変換、テストなど、ソフトウェア開発の様々なテーマのもとでの不確かさを扱った研究が存在する [1,4,6]。しかし、不確かさに関する実証研究はなく、そのため、実際にはどのような不確かさがどの程度存在するのか知られていない。そこで、我々は実際の OSS リポジトリを対象に、不確かさについての実証研究を行う。

不確かさはソフトウェア開発の各工程に存在している上に、対象となる問題の範囲が広すぎるため全体を捉えにくいものである。今回、本稿では revert に関する不確かさに着目して実証研究を行う。revert というのは、版管理システム Git の機能の 1 つで、過去のコミットの内容を打ち消すコミットをする機能である。revert の原因の 1 つとして、要求や仕様、実装手段などに不確かな内容があることで、後に変更されたり、必要がなくなるなどの理由で、過去の編集内容を打ち消す必要が発生するといったケースが存在する。そのため、revert と不確かさは関係性が強いのではないかと推測される。

本研究では revert コミットに注目し、実際のソフトウェア開発プロジェクトにおいてどのような不確かさが存在しているかを明らかにする。revert コミットには不確かさがどの程度の割合存在し、どのような不確かさが存在するのか目視調査で分類を行うことで、不確かさについて分析する。Git で管理されている 20 の OSS プロジェクトの開発履歴を対象に、以下の 2 つのリサーチクエスチョン (RQ) に答える。

RQ1: 不確かさを含む revert はどの程度存在するのか

まず、不確かさがどの程度影響をもたらす問題か理解するために、どの程度不確かさが存在しているか調査する。機械的に不確かさだと思われるコミットを検出し、集計する。全てのコミットに対する不確かさが占める割合と revert に着目した時に不確かさが存在する割合を比較することで、不確かさと revert にはどの程度関連があるかを確

¹ 九州大学

Kyushu University

a) muraoka@posl.ait.kyushu-u.ac.jp

b) kamei@posl.ait.kyushu-u.ac.jp

c) sato@posl.ait.kyushu-u.ac.jp

d) ubayashi@posl.ait.kyushu-u.ac.jp

かめる。

RQ2:どのような不確かさが revert には含まれるのか

機械的に発見された revert に関する不確かさを含むコミットの内容について、複数人で目視調査を行う。本研究では、まず Perez-Palacin らが定義しているソフトウェア工学における不確かさの分類の定義に基づいて、不確かさを分類する。また、既存研究での revert 理由に関する分類を用いて、revert に存在する不確かさの傾向について調査する。

以降、第2章では不確かさの分類に関する既存研究を紹介し、今回不確かさを分類する際に用いた定義を説明する。第3章では、不確かさの特定方法や目視調査の手法について説明し、第4章で結果について述べる。第5章では、今回の分析内容について議論を行い、第6章にて本研究のまとめとする。

2. 不確かさの分類

本章では、不確かさの分類に関わる既存研究を説明し、不確かさがどのように定義されているものであるか明らかにする。次に、不確かさの実証研究の問題について説明し、本研究において扱う不確かさの範囲について定義する。

2.1 ソフトウェア工学における不確かさの3つの分類

ソフトウェア工学において不確かさとは避けられないものである [2]。ソフトウェア開発における不確かさは大きく3つに分類される (*Known Knowns*, *Known Unknowns*, *Unknown Unknowns*)。 *Known Knowns* は不確かさが存在しない開発である。 *Known Unknowns* はソフトウェア開発のプロセスの中で不確かな問題が存在する開発で、開発者には何が不確かであるか認識されている状態である。 *Unknown Unknowns* は開発者が何が不確かであるか問題が理解できていない状態での開発を指している。

2.2 Perez-Palacin らによる不確かさの分類

Perez-Palacin らの研究 [4] では、ソフトウェア工学における不確かさを場所、レベル、性質の3つの要素で分類する方法を提案している (表1)。不確かさの場所とは、不確かさがソフトウェアのどこに現れたのかを指している。不確かさのレベルは、不確かさがどの程度開発者が問題に関する知識を持っているのかを段階的に分類したものである。また、不確かさの性質は不確かさが発生した状況についての分類である。

2.3 不確かさ実証分析における課題

不確かさとは曖昧なもので、不確かさの判断結果が主観によって大きく影響される可能性がある問題が存在する。Perez-Palacin らの研究の分類も、実際に分類するための基準が明確に設けられているわけではない。よって、同じ問

表1 Perez-Palacin らによる不確かさの分類

観点	性質	性質の説明
場所	コンテキスト	環境に関する不確かさ
	モデル構造	モデル自体の構造に現れる不確かさ
	入力パラメータ	モデルへの入力に関する不確かさ
レベル	レベル0	確定している知識
	レベル1	知識の不足を認知している状態。既知の不確かさ。
	レベル2	知識の不足を認知できていない状態。未知の不確かさ。
	レベル3	不確かさを認知するプロセス自体が不足している状態。
	レベル4	不確かさのレベル自体が不確か。
性質	認識的	十分なデータや知識が無いために発生する不確かさ
	偶発的	物理現象等の確率的不確かさ

題について不確かさの分類を行った時に、判断結果が人によって異なることもある。

本研究では、複数人の調査者でそれぞれ不確かさについて目視で評価し結果を集計することで、主観性の影響を小さくするようにした。詳しい分析の手法については3.3節にて説明する。

3. 調査手法

本章では、まず、不確かさをもつ revert コミットを特定する際に用いた、データセットとアプローチについて説明する。次に、特定したコミットに対して行った目視調査の手法について述べる。

3.1 データセット

本研究では Git で管理されている20のOSSリポジトリを対象とした。対象としたOSSのデータを表2に示す。revert コミットが存在するためには、ある程度の開発期間が必要だと予測される。これらのOSSは、開発期間も最短でも3年以上あり、様々な種類のOSSを用意したものである。

また、OSSリポジトリの情報のうちコミットメッセージを分析対象として扱った。開発履歴中の更新内容について記述するコミットメッセージには、更新内容に加えて、発生しているバグや不確かさなどの問題についても記述されている。

3.2 revert に関連する不確かさの特定方法

不確かさを持つ revert コミットを特定し分析する過程を図1に表す。

3.2.1 revert コミットの特定

revert コミットは、コミットメッセージの一部が定型文になっていて、revert されたコミットのコミットメッセー

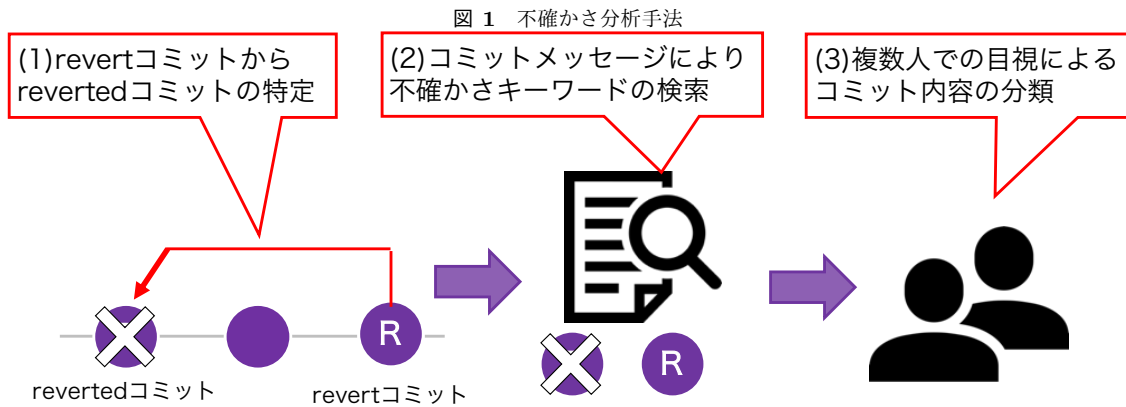


表 2 データセット概要

プロジェクト名	期間	全コミット数
apache/ant	2000/1/13-2017/12/12	13,710
apache/cassandra	2009/3/2-2017/11/21	23,687
apache/commons-codec	2003/4/25-2017/11/9	1,718
apache/commons-io	2002/1/25-2017/12/5	2,058
apache/commons-lang	2002/7/19-2017/12/20	5,128
apache/commons-logging	2001/8/2-2017/9/27	927
apache/derby	2004/8/11-2017/12/21	8,158
apache/james-mime4j	2005/5/3-2017/6/9	794
apache/tomcat	2006/3/27-2017/12/19	19,041
google/guava	2009/6/18-2017/12/20	4,520
mongodb/mongo	2007/10/19-2017/12/20	40,153
square/Aardvark	2014/10/5-2017/10/4	361
nicolasgramlich/AndEngine	2010/3/8-2013/12/11	18,005
android/frameworks-base	2008/10/21-2018/1/6	343,392
android/frameworks-opt-telephony	2012/6/13-2018/1/5	11,659
android/frameworks-support	2008/10/21-2017/12/7	35,641
android/tools-base	2012/12/10-2015/10/20	19,318
android/sdk	2008/10/21-2017/12/12	13,671
jedit/jedit-jEdit.bak	1998/9/27-2002/4/12	1,237
eclipse/eclipse.jdt.core	2001/6/5-2018/1/5	23,113
合計	—/—/— —/—/—	570,086

表 3 不確かさを表すキーワード

ambiguous	faltering	might	undecided
ambivalent	fluctuating	obscure	undetermined
arcane	fuzzy	of two minds	unforeseeable
blowing hot and cold	hesitant	open to question	unknown
chancy	in doubt	probably	unpredictable
changeable	in the balance	risky	unreliable
debatable	incalculable	tentative	unsettled
dicey	informal iffy	todo	unsure
doubtful	irregular	unascertainable	up in the air
dubious	irresolute	unclear	vacillating
erratic	may	unconfident	vague

ジの 1 行目とコミット ID についての情報が書かれている。これを利用して、データセットの中から、revert コミットと revert コミットで取り消されたコミット（以下、reverted コミット）のペアを検出する。

3.2.2 不確かさの特定

本研究では、コミットメッセージ内に不確かさのキーワードを含んでいるコミットを不確かさを含むコミットと定義する。本研究では不確かさを表すキーワードは、Uncertainty の類義語を *Oxford American writer's thesaurus* [3]

というシソーラス書籍から参照したものを用いた（表 3）。

また、revert コミットに不確かさがある場合と reverted コミットに不確かさが存在する場合の 2 種類に分けて分析を行う。これは、不確かさが revert に対してどう関わったが異なると予想されるためである。revert コミットに不確かさが存在する場合は、後から不確かさに関する問題が発覚したもの、reverted コミットに不確かさが存在する場合は、不確かさへの対応が適切でなかったために revert が発生したものではないかと推測できる。

3.3 目視での調査方法

機械的に検出した revert に関する不確かさを含むコミットについて目視で調査を行う。各不確かさに対し revert コミットと reverted コミット両方のコミットメッセージの内容を読み、内容について分類する。2.3 節で説明したように、不確かさの評価は主観による影響が大きい。そこで、本研究では 3 人の調査者でそれぞれ個別に分類を行い、以下の項目について不確かさを分類した結果を集計した。

- 不確かさの程度 - -2~2 の 5 段階
- 不確かさの場所 - コンテキスト, モデル構造, 入力パラメータ
- 不確かさのレベル - レベル 1~4 の 4 段階
- 不確かさの性質 - 認知的, 偶発的

まず、各 revert について不確かさの程度を -2~2 までの整数値を用いて 5 段階評価で分類する。不確かさの程度について、分類結果が 2 だとコミットには不確かさが含まれている、-2 はコミットには不確かさが含まれていないとし、0 は不確かさであるともないとも判断できないとして分類を行う。次に、不確かさの程度が 1 以上である（つまり、どちらかといえば不確かさである）と判断した revert は、2.2 節で説明した、Perez-Palacin らが定義した不確かさの分類の 3 要素（場所, レベル, 性質）についても不確かさを分類する。

4. 調査結果

本章では、1 章で挙げた 2 つの RQ について、3 章で説

表 4 不確かさ件数

project	全コミット数	不確かさを含む コミット数	割合 (%)	revert コミット数	割合 (%)	不確かさを含む revert コミット数	不確かさを含む reverted コミット
apache/ant	13,710	231	1.68	3	0.02	0	1
apache/cassandra	23,687	61	0.26	108	0.46	0	0
apache/commons-codec	1,718	16	0.93	0	0.00	0	0
apache/commons-io	2,058	22	1.07	2	0.10	0	0
apache/commons-lang	5,128	26	0.51	9	0.18	0	0
apache/commons-logging	927	15	1.62	0	0.00	0	0
apache/derby	8,158	247	3.03	0	0.00	0	0
apache/james-mime4j	794	7	0.88	1	0.13	0	0
apache/tomcat	19,041	284	1.49	0	0.00	0	0
google/guava	4,520	165	3.65	0	0.00	0	0
mongodb/mongo	40,153	318	0.79	856	2.13	2	6
square/Aardvark	361	1	0.28	1	0.28	0	0
nicolasgramlich/AndEngine	1,800	13	0.72	3	0.17	0	0
android/frameworks-base	343,392	3,587	1.04	1,777	0.52	20	55
android/frameworks-opt-telephony	11,659	165	1.42	68	0.58	0	1
android/frameworks-support	35,641	250	0.70	306	0.86	1	4
android/tools-base	19,318	138	0.71	35	0.18	0	0
android/sdk	13,671	166	1.21	23	0.17	0	1
jedit/jedit-jEdit.bak	1,237	10	0.81	0	0.00	0	0
eclipse/eclipse.jdt.core	23,113	70	0.30	22	0.10	0	0
合計	570,086	5,792	1.02	3,214	0.56	23	68

明した手法を用いて分析を行った結果について説明する。

4.1 RQ1:不確かさを持つ revert はどの程度存在するのか

分析結果を表 4 に示す。表中の割合は、プロジェクト全体のコミット数に対する不確かさを持つコミット数、revert コミット数の割合である。

全てのコミットの中に不確かさを含むコミットは 5,792 件 (全てのコミットの 1.02%)、revert コミットは 3,214 件 (全てのコミットの 0.56%) 存在した。revert コミットのうち、revert 時にコミットメッセージに自動で記入される reverted コミットのメッセージの 1 行目部分にキーワードを含むものは、重複を除くために、不確かさを含む revert コミットにはカウントせず、不確かさを含む reverted コミットとして集計とした。集計した結果、不確かさを含むコミット revert コミットは 23 件、reverted コミットは 68 件存在した。

今回対象とした OSS リポジトリ 20 個の合計では、不確かさを含む revert は、全ての revert コミットの 2.83% を占めていた。全てのコミットに含まれる不確かさコミットの割合 1.02% であるので、revert コミットには普段のコミットより高い割合で不確かさが含まれる結果となった。しかし、OSS プロジェクトごとに比較すると不確かさを含む revert、reverted コミットが 0 件のリポジトリが 14 個存在し、プロジェクト単位で比較すると revert コミットには高い割合で不確かさが存在するとは言えない OSS の方が多く存在した。

調査対象とした 20 リポジトリの全体のコミット 570,086 件中、不確かさを含むコミット revert コミットは 23 件、reverted コミットは 68 件存在した。

4.2 RQ2:どのような不確かさが revert には含まれるのか

RQ1 の機械的な分類で検出された不確かさを含む revert コミット 23 件すべて、reverted コミットは 68 件のうち 10 件を対象に調査を行った。分類結果を各分類の項目ごとに説明する。

不確かさの程度

不確かさの程度について分類した結果のうち、revert コミットに関するものを表 5 に示す。表中のキーワードは、コミットメッセージに含まれていた uncertainty の類義語である。また、A~C の列は 3 人の調査者の分類結果を表している。revert 理由については、島垣らの研究 [5] で用いられていた revert 理由に関する分類を利用し、コミット内容の補足とするために掲載した。

revert コミットのコミットメッセージに不確かさのキーワードが存在した 18 件については、3 人とも不確かさが存在すると判断したコミットが多い傾向にあった。全ての不確かさを含む revert コミット 23 件のうちに、例外は 2 件存在していた。2 件のコミットメッセージの一部を表 7 に示す。

1 件はコミットメッセージにエラーメッセージが貼り付けてあるコミットで、エラーメッセージ部分にある unknown

表 5 revert コミットに不確かさがある場合の不確かさの程度

プロジェクト	キーワード	revert コミット ID	reverted コミット ID	A	B	C	平均	中央	偏差
android/frameworks-base	may	9deba6	8ac32e	2	2	1	1.67	2.00	0.67
android/frameworks-base	risky	28f81d	d224c8	2	2	2	2.00	2.00	0.00
android/frameworks-base	may&probably	0f8ffd	b999cc	2	1	2	1.67	2.00	0.67
android/frameworks-base	probably	c82e7d	0cac71	2	2	2	2.00	2.00	0.00
android/frameworks-base	unclear	d6b42a	39aa79	2	2	2	2.00	2.00	0.00
android/frameworks-base	may	48b908	0a4e11	2	2	1	1.67	2.00	0.67
android/frameworks-base	probably	9309c1	—	2	2	2	2.00	2.00	0.00
android/frameworks-base	risky	75c21c	3127c2	2	2	2	2.00	2.00	0.00
android/frameworks-base	might	753e9e	c83d3e	2	2	2	2.00	2.00	0.00
android/frameworks-base	probably	4da482	059f8f	2	2	2	2.00	2.00	0.00
android/frameworks-base	unknown	ced57d	fe37ac	2	2	2	2.00	2.00	0.00
android/frameworks-base	unknown	ecd302	fe37ac	2	2	2	2.00	2.00	0.00
android/frameworks-base	probably	0b52ed	41bb49	2	2	2	2.00	2.00	0.00
android/frameworks-base	unknown	8b4838	77fe44	-1	0	-2	-1.00	-1.00	2.00
android/frameworks-base	might	7980d1	001d51	2	2	2	2.00	2.00	0.00
android/frameworks-base	probably	c5d8f6	27cf4a	2	1	2	1.67	2.00	0.67
android/frameworks-base	unknown	16626a	da3050	2	1	2	1.67	2.00	0.67
android/frameworks-base	probably	9b7a15	a47a9a	-2	1	1	0.00	1.00	6.00
android/frameworks-base	may	813805	1f5760	2	2	2	2.00	2.00	0.00
android/frameworks-base	might	b5428e	151819	2	1	0	1.00	1.00	2.00
android/frameworks-support	might	c56708	e5acd9	2	0	2	1.33	2.00	2.67
mongodb/mongo	may	7b9fe4	712c4f	2	1	2	1.67	2.00	0.67
mongodb/mongo	may	a7be5e	e6f9ef	2	2	1	1.67	2.00	0.67

表 6 reverted コミットに不確かさがある場合の不確かさの程度

プロジェクト	キーワード	revert コミット ID	reverted コミット ID	A	B	C	平均	中央	分散	revert 理由
android/frameworks-base	unknown	2ec157	3861bf	2	1	2	1.67	2	0.67	Defect introduced
android/frameworks-base	unknown	17548b	80ba0a	2	1	1	1.33	1	0.67	Obsolete solution
android/frameworks-base	unknown	3c9c09	17548b	2	1	1	1.33	1	0.67	Obsolete solution
android/frameworks-base	may	e4e7cd	a8a200	2	2	1	1.67	2	0.67	Obsolete solution
android/frameworks-base	unknown	e4e667	260c69	2	1	1	1.33	1	0.67	Obsolete solution
android/frameworks-base	unknown	2200da	993203	2	1	2	1.67	2	0.67	Incomplete fix
android/frameworks-base	may	4c5223	a8a200	2	2	1	1.67	2	0.67	Obsolete solution
apache/ant	may	296086	02aedb	2	0	2	1.33	2	2.67	Obsolete solution
mongodb/mongo	may	63d021	716e0f	2	2	2	2.00	2	0.00	Obsolete solution
mongodb/mongo	ambiguous	5c3f83	2eea3f	2	0	2	1.33	2	2.67	Obsolete solution

が不確かさのキーワードとして含まれているものであり、機械的には不確かさを含むと判定されたコミットであるが、調査者 3 人全てが不確かさだとは言えないと判断した。もう 1 件は、更新した内容がサイトに更新される日付が不確かだといった内容で、これをソフトウェア開発上の不確かさとするか、別の領域の不確かさとするかで意見が分かれていた。

コミットの内容としてはバグに関するものが多数を占めていた。表 5 の 23 件中 12 件がコミットメッセージ内にバグ ID を含んでいた。また、revert 理由はバグの原因が曖昧だからととりあえずバグを発生させたコミットを取り消したケースが多く見られた。具体例を 2 件、表 8 に示す。

不確かさの程度について分類した結果のうち、reverted コミットに関するものを表 6 に示す。

reverted コミットのコミットメッセージに不確かさが含まれる 10 件についても、不確かさでないと言われたコミットは存在しなかった。これらのコミットは、キーワードとして「Unknown」が用いられているコミットが多く、コミットの内容は、未知の入力や想定されていないイベントに対応するためのコミット内容が正しくなく revert されたケースが多い傾向にあった。unknown を含んでいた reverted コミットメッセージの一部を表 9 に示す。

表 7 例外の revert コミット

プロジェクト	コミット ID	revert コミットメッセージの一部
android/frameworks-base	8b4838	02-15 02:31:59.785 770 770 E System : Caused by: java.lang.IllegalArgumentException: Unknown component: com.android.bluetooth/com.android.bluetooth.opp.BluetoothOppLauncherActivity
android/frameworks-base	9b7a15	The reversion won't take effect on the site until the next scheduled DAC update, probably April 5.

表 8 バグと関連する revert コミット

プロジェクト	コミット ID	revert コミットメッセージの一部
android/frameworks-base	ecd302	Bug 6357880 has unknown causes at the moment but this change is the most likely candidate.
mongodb/mongo	7b9fe4	Unit test is failing on buildslaves; counts of records for indexes may be messing things up

表 9 unknown を含む reverted コミット

プロジェクト	コミット ID	revert コミットメッセージの一部
android/frameworks-base	17548b	Revert "be more robust with handling unknown sensors"
android/frameworks-base	e4e667	Revert "Treat UNKNOWN tootype events as Mouse events."

表 10 revert コミットに不確かさを含む場合の不確かさの場所

	場所	件数
完全一致	コンテキスト	1
	モデル構造	6
	入力パラメータ	0
部分一致	コンテキスト&モデル構造	15

数存在していた。

不確かさの場所について分類した結果のうち、reverted
コミットに関するものを表 11 に示す。

不確かさを含む reverted コミットでは、10 件のコミット
のうち調査者の意見が完全に一致するケースが 1 件のみで
あった。revert コミットでは見られなかった入力パラメー
タに存在する不確かさが多い傾向となった。

表 11 reverted コミットに不確かさを含む場合の不確かさの場所

	場所	件数
完全一致	モデル構造	1
部分一致	コンテキスト&モデル構造	1
	コンテキスト&入力パラメータ	3
	モデル構造&入力パラメータ	2
	各場所 1 件ずつ	3

revert コミットでは、22 件中 21 件のコミットがモデル
構造に不確かさが含まれていると判断する調査者が存在
した。また、入力パラメータに関する不確かさは存在
しなかった。reverted コミットでは、revert コミット
より不確かさの場所に関する分類結果が一致すること
が少なかった。

revert コミットでは、23 件中 21 件のコミットで不確か
さが存在すると判断され、バグに関するものが多かつ
た。reverted コミットでも、10 件中 10 件のコミットで
不確かさが存在すると判断され、例外対応されたコミ
ットが打ち消されたものが多く存在した。

不確かさの場所

不確かさの場所について分類した結果のうち、revert コ
ミットに関するものを表 10 に示す。完全一致は、不確か
さの程度 1 以上と分類した人全てが同じ不確かさの場所を
選んだコミットを表している。部分一致は、不確かさの程
度 1 以上と分類した人が異なる不確かさの場所をを選んだ
コミットを表している。

不確かさを含む revert コミットには入力パラメータに存
在するものはなく、コンテキストかモデル構造のどちらか
にしか不確かさが見られなかった。また、コンテキストと
モデル構造で意見が分かれたコミットが 22 件中 15 件と多

不確かさのレベル

不確かさの程度の調査で不確かさであると判断されたコ
ミット全てが、不確かさのレベルは 1 であると判定されて
いた。これは、コミットメッセージに不確かさのキーワ
ードを含むものを目視調査の対象としたことが原因であると
考えられる。対象コミットは、コミットに認識された内容
が記述されているコミットであるため、認知されていない
不確かさであるレベル 2 以上のものが存在しない結果に
なったと推測される。

発見された不確かさを含む revert コミット、reverted コ
ミットの不確かさのレベルは全てレベル 1 であった。

不確かさの性質

不確かさの性質について分類した結果を表 12 に示す。
不確かさの程度が 1 人以上の調査者に 1 以上と分類さ

表 12 不確かさを含むコミットの不確かさの性質

コミットの種類	不確かさの性質	件数
不確かさを含む revert コミット	認識的	21
	認識的&偶然的	1
不確かさを含む reverted コミット	認識的	9
	認識的&偶然的	1

れた 32 件のコミットのうち、30 件のコミットが調査者全てに認識的な不確かさであると分類された。残り 2 件のコミットについては、認識的と偶発的で調査者の意見が分かれた。

発見された不確かさを含む revert コミット、reverted コミットの不確かさの性質は 32 件中 30 件が全員一致で認識的な不確かさだと分類された。残りの 2 件は、認識的、偶然的で意見が分かれる結果となった。

5. 議論

今回の実証分析では、20 個の OSS を対象に分析対象とした。しかし、実際に目視調査を行った不確かさを含む revert コミット、reverted コミットは 4 つの OSS リポジトリにしか存在していなかった。また、OSS プロジェクトごとのコミット数をみると android/frameworks-base のコミット数が非常に多く、コミット数で見ると全体の約 6 割を占めている。さらに、目視調査を行ったコミット 33 件のうち、28 件が android/frameworks-base プロジェクトのものである。今回、revert コミットに存在する不確かさについて見つかった傾向が一般的なものでなく、android/frameworks-base プロジェクト固有の傾向である可能性が存在する。一般的な不確かさと revert についての実証分析を行うには、さらに分析対象リポジトリを増やす必要がある。

また、今回 4 人の調査者で分析を行なったが、4 章では、3 人分の結果しか表示していない。表に載せていない 1 人の調査者の結果が残りの 3 人と大きく異なるものであった。大きく差があったのは不確かさの程度の項目で、revert コミットについては他の 3 人の調査者と同様ほとんどのコミットを不確かさだと判断していた。しかし、reverted コミットでは、他の 3 人の調査者は不確かさではないと判断したものはなかったのに、表に載せていない調査者の結果では 10 件中 9 件のコミットに不確かさではないとする -2 をつけていた。不確かさとしなかった理由を表 9 の 1 つめのコミット (ID:17548b) の場合について確認した。ここで用いられている unknown は、もともとプログラムで対象としていなかったセンサーを指しているものであり、コミッタにとって未知のものではなかったのではないかと考え不確かさでないとしていた。一方、このコミットを不確かさと判断した調査者の理由は、以前は想定していな

かったセンサーの必要性に後から気づき変更したものであると判断し、Unknown Unknowns に相当する不確かさが known Unknowns に変化した不確かさだと捉えていた。

不確かさの調査を行う際、不確かさを分類するための情報が十分でない可能性があるといった問題が存在する。実証分析として、OSS の開発履歴から不確かさについて判断するが、開発履歴には開発時に開発者が考えていたことが全て記述されているとは限らない。今回、分析にはコミットメッセージを用いたが、コミットメッセージからコミッタの考えていることが一意に読み取れないケースも存在した。また、当然開発者が気づいていない問題についても記述されていない。つまり、2.1 節で説明した *Unknown Unknown* に値する不確かさについては、特定ができていない。今回の目視調査結果も、不確かさのレベルは全てのコミットがレベル 1、つまり、知識の不足を認識している状態、*Known Unknowns* の不確かさしか見つけることができなかった。

6. まとめ

本稿では、20 個の OSS リポジトリのコミットメッセージを分析対象として 2 の RQ を立て、実際のプロジェクトで発生する不確かさを含む revert について調査した。各 RQ において以下のような結果が得られた

(RQ1) 不確かさを含む revert はどの程度存在するのかという RQ の結果として、調査対象とした 20 リポジトリには、全体のコミット 570,086 件中不確かさを含むコミット revert コミットは 23 件、reverted コミットは 68 件存在し、revert コミットには、普段より高い割合で不確かさのキーワードが含まれていた。

(RQ2) どのような不確かさが revert に含まれるのかという RQ では、4 つの視点から不確かさについて 33 件のコミットを対象に 3 人で目視分析を行った結果を集計した。結果として、不確かさを含む revert コミット、reverted コミットともにレベル 1 の認識的な不確かさが多く存在していた。それぞれの特徴としては、不確かさを含む revert コミットには、ソフトウェアのモデル自体に不確かさが含まれてるコミットが多く、不確かさを含む reverted コミットには、revert 理由は誤った解決策を取り消したものが多く傾向にあった。

謝辞 本研究は、JP26240007 による助成を受けた。

参考文献

- [1] Elbaum, S. and Rosenblum, D. S.: Known Unknowns: Testing in the Presence of Uncertainty, *Proceedings of the 22nd International Symposium on Foundations of Software Engineering*, pp. 833–836 (2014).
- [2] Garlan, D.: Software engineering in an uncertain world, *Proceedings of the FSE/SDP workshop on Future of software engineering research*, ACM, pp. 125–128 (2010).

- [3] Lindberg, C. A.(ed.): *Oxford american writer's thesaurus*, Oxford University Press (2012).
- [4] Perez-Palacin, D. and Mirandola, R.: Uncertainties in the Modeling of Self-adaptive Systems: A Taxonomy and an Example of Availability Evaluation, *Proceedings of the 5th ACM/SPEC International Conference on Performance Engineering*, pp. 3–14 (2014).
- [5] Shimagaki, J., Kamei, Y., McIntosh, S., Pursehouse, D. and Ubayashi, N.: Why are commits being reverted?: a comparative study of industrial and open source projects, *Software Maintenance and Evolution (ICSME), 2016 IEEE International Conference on*, IEEE, pp. 301–311 (2016).
- [6] Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B., Janssen, P. and Krayen von Krauss, M. P.: Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support, *Integrated assessment*, Vol. 4, No. 1, pp. 5–17 (2003).