

# OSS プロジェクトにおける不確かさの内容分類

村岡 北斗<sup>†</sup> 亀井 靖高<sup>†</sup> 佐藤 亮介<sup>†</sup> 鷗林 尚靖<sup>†</sup>

<sup>†</sup>九州大学

## 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、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く、バグやソースコードの煩雑化の原因となりがちな問題である。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 [1]。

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。しかし、不確かさに関する実証研究はなく、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 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。

本稿では不確かさの内容について調査を行うこととした。具体的には、開発者が開発中のどのような出来事について不確かであるのか、つまり何について開発者の知識が足りない状態であるかといった観点で分類を行なった。不確かさの分類については、Perez-Palacin らの研究 [2] が存在しているが、これは不確かさの性質的な面について分類を行なったものであり、今回の内容とは異なるものである。実際のソフトウェア開発においてどのような不確かさに開発者が困っているか内容を分類することで、どのような不確かさが存在していて、どの不確かさが重要な問題なのか知ることができ、不確かさをうまく管理したソフトウェア開発について考えることができる。

また、本研究では revert に関わった不確かさと、そ

うではない不確かさに対象を分けることで開発に悪影響をもたらす不確かさに傾向があるのではないかと調査を行なった。revert とは、git の機能の 1 種であり、過去に行なった更新を取り消すものである。revert に関連した不確かさを、不確かさが原因で必要なかったはずの工程が発生してしまったものではないかと仮定し、revert と関連していないコミットの内容と比較することで擬似的に開発に悪影響をもたらす不確かさに傾向があるのか調査した。

## 2 調査手法

### 2.1 データセット

本研究では Git で管理されている 20 個の OSS プロジェクトを対象とした。これらの OSS プロジェクトはコードスメルに関する実証分析の論文でもデータセットとして扱われていたものである。revert コミットが存在するソフトウェア開発には、ある程度の開発期間が必要だと予測される。これらの OSS プロジェクトは、開発期間も最短でも 3 年以上あり、また様々な種類の 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。今回は、20 プロジェクトの全コミット合計 570,086 件を対象とした。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。

### 2.2 不確かさの特定

データセットの 20 プロジェクトの全てのコミットを対象に、コミットメッセージに Uncertainty の類義語が含まれているものを、不確かさ候補コミットとして扱った。また、コミットコメントの内容を利用して、不確かさ候補コミットが revert コミットか revert によって打ち消されたコミットでないかについても調査した。特定された不確かさ候補コミットについて表 1 に示す。

### 2.3 不確かさの分類

2.2 節で発見された不確かさ候補コミットを対象にオープンコーディングを用いて、不確かさの内容につ

### Classification of Uncertainty in OSS Projects

Hokuto Muraoka<sup>†</sup>, Yasutaka Kamei<sup>†</sup>, Ryosuke Sato<sup>†</sup> and Naoyasu Ubayashi<sup>†</sup>

<sup>†</sup>Kyushu University, 744 Motoooka, Nishi-ku, Fukuoka, Japan

<sup>†</sup>muraoka@posl.ait.kyushu-u.ac.jp

{kamei, sato, ubayashi}@ait.kyushu-u.ac.jp

表 1: 見つかった不確かさ候補コミット数内訳

	コミット数	割合
不確かさ候補コミット	5,792	100.0%
revert コミット	80	1.38%
非 revert コミット	5,712	98.62%

表 2: 不確かさ分類結果

カテゴリ	revert なし	revert あり
例外的な内容の不確かさ	13	6
バグと関連した不確かさ	29	21
将来的な変更についての不確かさ	3	7
変更内容に対する不確かさ	4	6
コードの内容に関わる不確かさ	4	0
その他不確かさ	9	10
不確かではない	4	5
合計	100	80

いて分類した。revert に関連したコミットは 80 件全て、そうでないコミットは 5,712 件中からランダムに 100 件を選び対象とした。今回 1 人の調査者で各不確かさ候補コミットについて目視調査を行なった。まず、コミットコメントの内容を簡単な文章で説明したものに交換した。次に似た内容のものについて組み合わせ、各集合に名前をつけることでカテゴリとした。また、この過程で不確かさ候補コミットに記述されている内容が不確かさかどうかの判定も行なった。

### 3 調査結果

分類された不確かさ候補コミットのカテゴリごとの件数を表 2 に示す。コミットメッセージに現れる不確かさは 5 種類のカテゴリに分類することができた。

- ・例外的な内容の不確かさ: 未知の入力への対応や入力に対する結果が曖昧な不確かさ。
- ・バグと関連した不確かさ: バグの原因、影響、修正に関係した内容の不確かさ。
- ・将来的な変更についての不確かさ: 将来的に改善や修正のために違うアプローチによる実装が必要とされるかもしれないという内容の不確かさ。
- ・変更内容に対する不確かさ: そのコミットで行なった変更内容が不十分であるかもしれない、新たな問題を生み出してしまったかもしれないといった内容の不確かさ。
- ・コードの内容に関わる不確かさ: コミッタがプログラムの振る舞いが理解できていないため、変更が正しいかどうか判断ができない不確かさ。

今回調査したデータセットでは、revert の有無に関わ

らずバグに関連した内容の不確かさが多く観測された。

### 4 議論

今回、不確かさの目視分類を 1 人で行なったため、主観による影響が強い結果になっている可能性がある。そのため、複数人で不確かさの内容について議論する必要が必要であると考えられる。

また、今回不確かさを 5 つのカテゴリに分類する結果になったが、カテゴリの内容が時間による分類（将来的な変更についての不確かさ、現在の変更内容に対する不確かさ）と何が不確かなのか（例外、バグ、コードの内容）で分類したもの、2 種類が混合して存在してしまっている。この問題については、時間による分類と何が不確かであるかの分類、それぞれの観点で個別に再分類を行なってみるなどの対策が考えられる。

### 5 おわりに

本稿では、20 個の OSS リポジトリを対象に不確かさの内容について調査した。コミットメッセージについて目視で分類を行なった結果、コミットメッセージに現れる不確かさは 5 種類に分類することができた。今回のデータセットでは、バグに関わる内容の不確かさが観測された不確かさ全体の約 4 割と多く観測される結果となった。

今回の分析は不確かさの内容について行なった実証研究の第一歩である。将来的に、さらに不確かさについての分析を重ねることで、開発者に必要とされている不確かさを考慮した開発手法について考えていく予定である。

#### 謝辞

本研究は、JP26240007 による助成を受けた。

#### 参考文献

- [1] 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).
- [2] 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).