

# ソースコードに存在する不適切なコメントの検出手法適用事例

甲斐 秀一<sup>1)</sup>      田上 諭<sup>2)</sup>      片山 徹郎<sup>3)</sup>

1) 株式会社スカイコム h-kai@skycom.jp

2) 3) 宮崎大学 工学研究科 {tanoue, kat}@earth.cs.miyazaki-u.ac.jp

## 要旨

本研究では、ソースコード(最終的にはソフトウェア)の品質向上を目的とし、ソースコード内に存在する不適切なコメントを検出可能なツール(以下、本ツール)の開発および評価を、宮崎大学と共同で行った。不適切なコメントとして検出した箇所のソースコードを見直すことによって、品質およびメンテナンス時の作業効率の向上を図ると共に、不適切なコメントが及ぼす影響やコメントの重要性を理解するための教育ツールとしても利用可能なものを目指して取り組みを行った。

## 1. はじめに

ソースコードに用いられるコメントは、プログラムの処理には直接影響を与えないため、内容について重要視されにくい傾向にある。当社でのソースコードレビュー時においても、コードに注視してチェックを行うため、コメント内容の細かいチェックまでは実施しない傾向があった。そのため、不適切なコメントを含んだソースコードが、知らぬ間に徐々に蓄積されていた。読み手にとって理解の妨げとなるような不適切なコメントが存在する場合、メンテナンス時の作業効率を悪化させ、バグの混入に繋がり、結果として製品の品質を低下させる可能性が考えられる。

そこで本研究では、ソースコードの品質向上を目的として、不適切なコメントを容易に検出する仕組みを適用した場合の効果を検証するため、本ツールの開発および評価を行った。

## 2. 不適切なコメントの検出手法

検出すべき不適切なコメントについては、当社開発部門内でアンケート調査を実施して定義した。アンケート調査の結果、社員が不適切と考えるコメントのほとんどが、単語によってパターン分けできることが判明した。(詳細はスライド参照)

以下に検出したパターン例を挙げる。

- 意図が曖昧なコメント  
(例)「暫定」「仮」「未定」

- 周辺処理を理解せず追加したと推測されるコメント  
(例)「なぜか」「おそらく」「かもしれない」

前述のアンケート結果をもとに、不適切になり得る単語を一覧化した定義ファイルを用意し、定義した単語を全コメント中から検出することで不適切なコメントの検出を行うこととした。なお、定義する文字列には正規表現も使用可能にしているため、検出文字列のパターンは柔軟に対応可能である(例えば、日付フォーマットを含むコメントの検出等も可能)。

## 3. 本ツールの評価

当社開発部門が保有する複数ソースコードに本ツールを適用した結果、主に以下の効果が得られた。

- 不適切なコメントの目視確認が不要となり、過去の実装も含めて容易に見直しが可能となった。
- 検出対象コメントの定義を柔軟に拡張できるため、ツールを実行して検出結果の傾向を分析し、定義ファイルを拡張していくことで検出精度が上がった。
- 検出した不適切なコメントの位置から、その周辺処理に潜在する課題や見直しが必要な可能性のある処理の検出に成功した。

(詳細はスライド参照)

## 4. おわりに

不適切なコメントを検出し、その周辺箇所を見直すことによって、品質およびメンテナンス時の作業効率の向上に繋がることを確認でき、改めてコメントの重要性を認識した。また、本ツールを利用することによって、効率良く不適切箇所を検出できることを確認した。

今後は、不適切なコメントだけでなく、コードとコメントが一致していない箇所の検出を可能とするなど本ツールを拡張することによって、さらなる開発全体の作業効率化を促進したい。