

品質属性の分類と指摘文の類型化を用いた ソフトウェア開発文書診断法

藤田 悠[†] 山本 雅基^{†,†††} 中澤 達夫^{†,††††} 塩谷 敦子^{††}
池田 貴一^{†††} 楡井 雅巳[†] 小野 伸幸[†]

我々は、企業の現場で作成される開発文書のレビューを行い、要求分析や設計などの開発技術上の指摘だけでなく、開発文書としての役割を果たしていない点や、複数の解釈を許すようなあいまいな記述などを指摘してきた。しかし、問題点を指摘する表現と分類がレビューに依存することに直面した。

これまで我々が行った文書診断法では、文書を読んで受ける印象や直し方の示唆など、問題点を指摘するために多様な表現が用いられていた。さらに、様々な問題点を分類する基準が明確でなかった。そこで、まず、技術文書が備えるべき品質属性を与え、次に、問題点を指摘する文を類型化して要約文を作成した。この手法を用いることで、品質属性は問題となる原因に対応づけられることが明らかになった。

Document Diagnostic Method by Classification of Document Quality Characteristic and Typology of Explanatory Sentence

YUTAKA FUJITA,[†] MASAKI YAMAMOTO,^{†,†††} TATSUO NAKAZAWA,^{†,††††} ATSUKO SHIOYA,^{††}
KIICHI IKEDA,^{†††} MASAMI NIREI[†] and NOBUYUKI ONO[†]

We reviewed software development documents produced by company development activities. We checked defects of not only software development engineering about requirement and architecture but also description as a development document and technical writing with ambiguous representation. Then we confronted with problems that representation and classification depend on reviewers.

Previously we have pointed out defects on development documents using description of impression or revision proposal and we haven't defined quality characteristic of document.

At first we defined quality characteristic about technical document. Next, we created abridged sentence that represented defects corresponding to quality characteristic. In the result, each quality characteristic specified with cause of problems respectively.

1. はじめに

プログラムには、コードの書き方の規則を与える MISRA-C[1] や Indian Hill C Style and Coding Standards[2]などのコーディング規約や、組込みソフトウ

ェア向けコーディング作法ガイド[3]などのコーディング作法が与えられている。他方、ソフトウェア開発文書には標準的な記述の規則や指針は見当たらない。もちろん、例えば、要求定義書について、IEEE830-1998[4]はソフトウェア要求定義書が満たすべき性質として8種類の性質(正当性、無曖昧性、完全性、一貫性、順位付け、検証容易性、修正容易性、追跡性)[5]を与えている。しかし、どのように書くことでこれらの性質が満たされるかについては示されていない。

日本語の記述のあいまいさに関しては、ソフトウェア開発文書[6][7]にとどまらず、様々な種類の文書の問題として提起されている[8][9][10][11]。これらが提起している日本語の表現における問題を理解することは、よ

[†] 長野工業高等専門学校
Nagano National College of Technology
^{††} 合同会社 イオタクラフト
IOTACRAFT,LLC.
^{†††} 株式会社 ミマキエンジニアリング
MIMAKI ENGINEERING Co.,Ltd.
^{††††} 名古屋大学
Nagoya University
^{†††††} 信州大学
Shinshu University

りよい文章を書くために有効であると、我々は捉えている。そこで、これらの知見を活用することにより、ソフトウェア開発文書の品質を向上させることが可能であると考えた。

我々はこれまで、実際にソフトウェア開発で書かれた文書に含まれる問題を分析し、その分析結果を手がかりにして技術者を教育するプログラムを立案し、実施してきた[12]。教育プログラムの立案は、開発文書をレビューして問題の指摘と分析を行い、その文書を書いた技術者や文書を活用する組織などが抱える課題を洗い出すことから始まる。そこで洗い出されたソフトウェア開発の課題や開発文書の記述に関する問題に対応する、題材を設定し、技術者が必要とする能力の向上や組織における課題を解決に導くための教育を行った。

我々は、開発文書を精査するレビューを文書診断、さらにレビューを診断者と呼び、レビューを新しい枠組みで捉えなおそうとしている。これまでに行った文書診断の課題として、文書に含まれる問題点の表現と分類が診断者に依存することがある。つまり、診断者が開発文書の内容を読んで診断するとき、書かれている文章から受けた印象や、直し方の示唆などの多種多様な表し方が用いられていた。さらに、文書の品質属性を定めていなかったため、傾向を分析するための分類の基準が明確になっていなかった。

そこで本稿では、初めに、これまでに行った文書診断の経験から、技術文書の品質属性を与える。次に、その品質属性に分類するために、問題がある表現の名称とその表現の状態を表す形式に、問題点を指摘する文章を類型化した。これら 2 種類の提案を文書診断法に適用し、各品質属性に分類される問題点の特徴を明らかにする。

2. 文書品質に関する先行研究

2.1. トップダウンによる品質分類

竹内ら[13]は、IEEE830-1998[4]が示す、要求仕様が満たすべき 8 項目(合目的性、非曖昧性、完全性、無矛盾性、重要度、順位付け、検証可能性、追跡可能性)に着目し、ソフトウェア開発文書の書き方における問題点の検出を、自然言語処理技術を用いて行っている。なお、自然言語処理技術における制約から、非曖昧性、完全性、検証可能性、追跡可能性の 4 項目のみに限定している。

Arthur ら[14]は、ソフトウェア開発文書が持つべき 4 種類の文書品質(正確性、完全性、使用性、拡張性)をもとに 13 種類の品質の要因に分解し、30 種類の数

量詞、37 種類の文書品質指標に分解して系統立てた分類木を与えている。Arthur らは、文書品質を計測するための枠組みとしてこの分類木を与えており、その検証や評価の方法については今後の検討課題であると述べている。

抽象的な属性をもとにして開発文書から問題点を指摘するときは、もともになる属性を詳細な項目に具体化することと、文書上の問題点を抽出する手法を与えることが必要であると考えられる。そうすると、具体化する項目の設定や与えた手法によって異なる結果が導かれると予想する。

2.2. 文書診断法

我々は、文書による情報伝達に支障を来す問題点を、文書の内容を読んで抽出する、文書診断法を提案した[15]。文章の表記だけでなく、内容に踏み込んで指摘を行う文書診断法では、文書診断を行う診断者が 2 種類の診断観点にしたがって問題点を抽出する。その 2 種類の診断観点は、技術文書に共通する日本語の文法や論理的な展開に関する「技術文書観点」と、ソフトウェア開発固有の開発プロセスとの関連や開発プロセスの入出力などに関する「開発文書観点」である。問題点を検出した際には、その内容を説明する指摘文を作成する。

指摘文を作成したあとに、5 種類の診断種別に関連付けられた 73 項目の診断項目に分類して、診断結果を集計する。まず、指摘した問題を分類するために、種類別に次の 5 種類に分類した。その診断種別は、文書体系と文章構造・構成・展開と文・文章の表現と図表の表現と表記ルールである。さらに、これらの診断種別それぞれに含まれる指摘の、大まかな内容が把握できるように 73 項目の診断項目に分類した。各診断種別に

表1 診断種別ごとの診断項目の例

診断種別	診断項目
文書体系	上位文書の存在が不明、文書の役割が不明確など 4 項目
文章構造・構成・展開	開発文書としての文書タイトルの付け方が不適切、タイトルと内容の不一致など 23 項目
文・文章の表現	根拠の記述なし、一文一義でないなど 22 項目
図表の表現	開発文書としての図表の基本的な表記ルール違反、図表の凡例の不備など 12 項目
表記ルール	改訂履歴の記載の不備、用字・用語の不統一など 12 項目

該当する診断項目を表1に示す。

2.3. 文書診断法の課題

従来の文書診断法では、診断者の印象などを含む表現で表されている指摘文や診断項目がある。これまでの診断では、特定の診断者が独自のとらえかたで項目の分類や結果の分析を行っていたので、その診断者が一貫して行っているときはよいが、文書診断法をほかの人が適用する場合や、診断と分析を異なる人が行う場合などには、問題を説明する表現をそろえる必要がある。さらに、種類別の分類ではそれぞれの問題が文書にどのような影響を及ぼすか分析できず、その傾向を判断する基準がなかった。

3. 文書診断法の改良

問題点の表現と分類が診断者に依存するという、文書診断法における課題を解決するために、改良した文書診断法を提案する。その提案には、文書が備えるべき品質属性に分類する手順と、品質属性に合わせて問題点を説明する文章を類型化して記述しなおす手順を追加した。これまでの文書診断で行ってきた、診断種別に関連付けられた診断項目に分類する手順に、新たな方法を加えた文書診断の手順を、フローチャートを用いて図1に示す。

3.1. 技術文書の品質属性

技術文書の品質属性として、「合目的性」と「論理性」と「表記の適切性」の3種類の品質属性を提案する。以上の品質属性は、これまでに行った文書診断の経

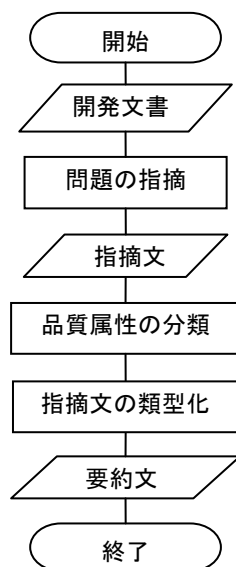


図1 品質属性の分類と指摘文の類型化を含む診断手順

験から導いた。

「合目的性」は、文書の役割を果たす記述がなされているか、文書の内容がその主旨に適合しているかなど、文書の目的に関する属性である。具体的には、ソフトウェア開発文書であれば、開発プロセスにおける作業工程の成果物として適切な内容が記述されているか、開発プロセスを意識した記述が表されているか、ソフトウェア開発にかかわる作業を行うために十分な項目が記載されているかなどである。

「論理性」は、技術的な内容まで立ち入らない範囲の、論理展開に関する属性である。具体的には、章や節や段落の構造や構成、文と文のつながり、接続詞の使い方などである。

「表記の適切性」は、日本語の文法や文章の体裁などの、内容に入り込まない範囲における表記に関する属性である。具体的には、文法や文書の形式や記号の使い方や言葉使いが適切であるか、などである。

これら3属性は、図2に示した階層構造で表される。まず、文書の最も基礎的な部分である「表記の適切性」を最下部に位置づける。次に、「論理性」を満たしているかを判断するには、「表記の適切性」を満たした文章を精査すると考えたので、第2層に「論理性」を位置する。さらに、「合目的性」を満たしているかを判断するには、「論理性」を満たした文章を精査すると考えたので、第3層に「合目的性」を位置する。

3.2. 指摘の要約文

文書診断で作成する指摘文から、表現をそろえて要約文を作るために、要約文に条件を与える。その条件は、要約文が、問題の原因となる表現の名称を表す「要素」を主部に持つことと、その「要素」の様相を表す「状態」を述部に持つことである。すなわち、主部と述部の間に助詞「が」を用いて接続することで、事物の現象を表す文章になる[16]。

「要素」と「状態」を表す語句の例を示す。「インクがどんどん流れる」という仕様が書かれていた場合、擬態語を使用することによってあいまいさが引き起こされていることを、診断者は指摘する。このことを表現する要約文の「要素」は、「どんどん」という擬態語そのものではなく、表現の名称である「擬態語」とする。他方、「状

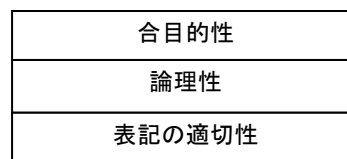


図2 技術文書の品質属性を構成する階層構造

表2 問題を含む文例とその問題を指摘する指摘文と要約文

	文例	指摘文	要約文
(1)	A に対して指示を出す前に、該当モータの出力パルス数を読み出す。 【理由】出力パルス数を読み出すことにより、FPGA 内部でカウンタがリセットされる。	カウンタをリセットするのが目的なのか、出力パルスを読み込むことが目的なのかわからない。	作業の目的が記述されていない
(2)	【仕様】モータの回転時間をタイマーで管理しない。 【理由】処理時間は、処理する対象によって異なるため。 【説明】動作中はポンプモータを回し続ける。	回し続けることが実行内容であつて、「タイマーで管理しない」というのは、それを実現するための補足情報であるので、関係が逆転している。	主題と補足の関係が逆転している
(3)	初期化時にタンクのパルブを閉めたまま動作を行いヘッドとタンクの間に負荷をかける。	「動作を行い」と「ヘッド」の語句がつながっている。	読点が正しく使われていない

態」を表す語句は「用いられている」である。

業の目的が記述されていない」という要約文を与えた。したがって、「要素」は「作業の目的」であり、「状態」は「記述されていない」である。

4. 改良した文書診断法の適用

診断者 2 名が全 6 ページからなる、大型プリンタ用の組込みソフトウェアの要求仕様書に対して文書診断を行った。そこで指摘した 136 件の問題点を品質属性に分類し、そこで与えた指摘文から要約文を作成した。

4.1. 品質属性の分類と要約文の作成と例

問題を含む文章とそれを指摘する指摘文と要約文の例を表 2 に示す。ただし、文例は製品の開発情報を秘匿するために、実開発文書の文言を一部改編している。なお、文例中の括弧(【,】)はすべて文中で用いられていたものであり、我々の注釈などではない。

文例(1)は、操作の目的に関する指摘であるので、品質属性を「合目的性」とした。ここで、パルス数を読み出す目的を文書から読み取ることができないので、「作

文例(2)は、文章の論理的な構成に関する指摘であるので、品質属性を「論理性」とした。ここで、「【仕様】」と「【説明】」に書かれている内容をみると、仕様と説明に合致する内容になっておらず、行うこととその理由の関係が逆になっている。そのため、「主題と補足の関係が逆転している」という要約文を与えた。「要素」は「主題と補足の関連」であり、「状態」は「逆転している」である。

文例(3)は日本語の表記にかかわる指摘であるので、品質属性を「表記の適切性」とした。ここで、「動作を行い」と「ヘッド」の間に読点が欠落しているので、「読点が正しく使われていない」という要約文を与えた。「要素」は「読点」であり、「状態」は「正しく使われていない」である。

4.2. 品質属性ごとの指摘件数

品質属性の傾向を示す。全指摘件数 136 件のうち「合目的性」が 24 件(18%)、「論理性」が 52 件(38%)、「表記の適切性」が 60 件(44%)を占める診断結果となった。品質属性ごとの指摘件数の割合を図 3 に示す。この結果は、この要求仕様書には文法や文章の体裁などの「表記の適切性」にかかわる問題と、文章のつながりや構成などの「論理性」にかかわる問題が多く含まれていることを示している。したがって、この文章では、基礎的な記述が十分でないという診断を下す。

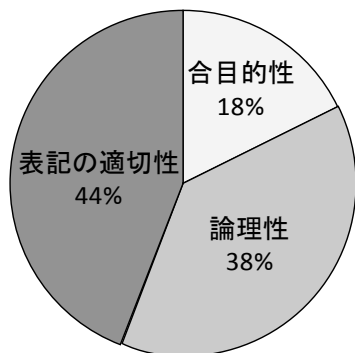


図3 品質属性ごとの指摘件数の割合

5. 考察

5.1. 要素の語句と品質属性の関連

本稿で行った文書診断で作成した要約文に含まれ

表3 各品質属性に関連付けた要約文の要素と状態を表す語句

	合目的性	論理性	表記の適切性
要素	未定義事項の解決を誰がいつ行うか、参照の目的や方法、作業の目的、意図している内容、など	主題と補足の関係、構成ごとに記述する内容、選択肢の定義、要求の説明、要求の理由、など	余計な空欄、用語定義、用語、主語、口語表現、助詞の使い方、など
	計 11 種類	計 29 種類	計 20 種類
状態	書かれていない、定まらない、不明、わからない	書かれていない、使われている、不明、持って回っている、明確でない、など	ある、誤解を生じる、省略されている、同じ対象に用いられている、など
	計 4 種類	計 23 種類	計 22 種類

る「要素」と「状態」を表す語句の一覧を表 3 に示す。

「合目的性」には、11 種類の「要素」を表す語句と 4 種類の「状態」を表す語句が表れた。同様に、「論理性」には 29 種類の「要素」と 23 種類の「状態」、「表記の適切性」には 20 種類の「要素」と 22 種類の「状態」を表す語句が表れた。

「合目的性」に分類した要約文の「要素」には「内容」や「目的」や「方法」などを含む語句が表れた。このことから、技術文書は、

- 文書の目的をかなえるために記述すべき内容
- 文書の位置づけ
- 文書を用いて行うべきこと

などにかかわる「要素」で具体化された「合目的性」を備えるべきであることを示唆していると考ええる。

「論理性」には、「関係」や「構成」や「項目」などを含む語句が表れた。このことから、

- 文書の目的に適した文章を構成する
- 内容の展開に飛躍が生じないように項目を立てる
- 章や段落や文を順序立てて展開する

などの論理展開に関連する、「要素」で具体化された「論理性」を備えるべきであることを示唆していると考ええる。

「表記の適切性」には、「用語」や「主語」や「動詞」や「長さ」などを含む語句が表れた。このことから、文法や文書形式などの規則に適合させることや、誤解を招かないように記述するなど、内容には深く入り込まない範囲にある「要素」で具体化された「表記の適切性」を備えるべきであることを示唆していると考ええる。

要約文の「要素」に含まれる語句には、「合目的性」と「論理性」と「表記の適切性」それぞれの品質属性に関連する語句が表れている。したがって、要約文の「要素」は問題点と品質属性を具体的に関連付ける役割を果たすと考える。このことから、品質属性に分類するときには、問題を引き起こしている表現の名称である「要素」を基準にすることで診断者に依存しない結果が得

られると考える。

5.2. 状態の語句と品質属性の関連

「合目的性」と「論理性」と「表記の適切性」のそれぞれに当てはまる要約文の「状態」を表す語句には、品質属性の違いにおける顕著な違いはみられなかった。「書かれていない」や「不明」など 3 属性に重複して表れている語句もあった。したがって、状態を表す語句には品質属性との関係は認められない。

しかし、指摘された問題を解消させるためには、その表現の「状態」を解消する必要があるため、「状態」は、それぞれに応じた対処方法を与えるための参考になると考える。今後は「状態」に応じた問題の対処方法を示すために、要約文の「状態」を表す語句をより深く観察する必要がある。

5.3. 文書診断における要約文の活用

要約文の「要素」には、文書を構成する部品や項目や表現法などの名称が用いられているので、診断者が何をどのように捉えて判断したのか明らかである。したがって、指摘文を類型化することによって、診断者の感覚が具体化されると考える。このことから、具体化された項目の一覧は、診断者のスキル測定や向上させるべきスキルを判断する基準になると考える。今後は、それぞれの品質属性を特徴づける「要素」に該当する語句の抽出と分類が必要である。

本稿では、3 種類の品質属性ごとに指摘した問題点の件数から、文書の品質を分析している。しかし、品質属性ごとの割合の傾向だけでなく、全体の指摘件数やその他の計測方法によって導かれる定量的な値をふまえて分析する必要があると考える。例えば、指摘件数が多い場合は、全般的な品質の向上を重視すべきである。他方、指摘件数が少ない場合は、品質属性の傾向を詳細に分析して対策を講じる必要がある。そのためには、ページあたりの指摘件数の基準値や、品詞ごとの語句の数などによる定量的に計数可能な値をふまえて分析することが有効であると考えている。

6. まとめ

文書の問題点を指摘する文書診断法では、問題点の表現と分類が診断者に依存していた。

そこで、技術文書が備えるべき3属性からなる品質属性として「合目的性」と「論理性」と「表記の適切性」を与えて、これらに分類するために、問題の原因となる「要素」と「状態」を表すように指摘文を類型化した要約文を作成する方法を提案した。この手法を用いることで、品質属性に分類され、問題点を表す表現がそろっているので、診断者への依存が抑えられると考えた。

品質属性への分類と類型化した要約文の作成、これら2種類の提案を用いて文書診断を行ったところ、今回の診断対象とした実開発文書では「論理性」と「表記の適切性」の指摘が多く検出され、基礎的な文書の記述が十分でないことが明らかになった。

要約文の「要素」と「状態」を表す語句を分析したところ、「要素」を表す語句には品質属性ごとに特徴が表れたが、「状態」を表す語句には品質属性の間での顕著な違いが認められなかった。このことから、品質属性ごとの特徴は、要約文の「要素」に現れることを確認した。

指摘文の類型化によって、問題点を指摘するときの診断者によるとらえかたを、具体化できると考える。それによって得られた結果は、診断者のスキルの測定や向上させるべきスキルを判断する基準になると考える。

今後は、原因を表す「要素」とその様相を表す「状態」に該当する語句の蓄積と、それを手がかりにした文書診断者スキルの評価や向上の方法の検討、さらに定量的な計測可能な値とその標準的な値をふまえて診断結果を分析する手法の検討を行う。

参考文献

- [1] 日本規格協会, MISRA-C 研究会: 組込み開発者における MISRA-C:2004-C 言語利用の高信頼化ガイド, 2006
- [2] L. W. Cannon and R. A. Elliott and L. W. Kirchhoff and J. H. Miller and J. M. Milner and R. W. Mitze and E. P. Schan and N. O. Whittington and Henry Spencer and L. W. Cannon and R. A. Elliott and L. W. Kirchhoff and J. H. Miller and J. M. Milner and R. W. Mitze and E. P. Schan and N. O. Whittington and Henry Spencer: Indian hill c style and coding standard , 1997
- [3] 経済産業省, (独)情報処理推進機構 ソフトウェア・エンジニアリング・センター: 改訂版 組込みソフトウェア開発向け コーディング作法ガイド[C言語版], 2007
- [4] IEEE Computer Society: IEEE Recommended Practice for Software Requirements Specifications, IEEE Std 830-1998, 1998
- [5] 山本修一郎: 要求を可視化するための要求定義・要求仕様書の作り方, ソフトリサーチセンター, 2006
- [6] 塩谷敦子, 菊池百代: ソフトウェアドキュメンテーション, 株式会社デンソークリエイト, 2000
- [7] 福田 修: SEを極める 仕事に役立つ文章作成術—百戦錬磨のプロマネが伝授するドキュメント作成の極意, 日経 BP 社, 2005
- [8] 風間 力三: 文章ドクター 悪文の診断と治療, 東京堂, 1962
- [9] 岩淵 悦太郎: 悪文 第3版, 日本評論社, 1989
- [10] 木下 是雄: 理科系の作文技術, 中央公論新社, 1981
- [11] 阿部 圭一: 明文術 伝わる日本語の書きかた, NTT 出版, 2006
- [12] 藤田 悠, 山本 雅基, 中澤 達夫, 塩谷 敦子, 楡井雅巳, 小野 伸幸: 長野高専寄附研究部門による企業のソフトウェア開発文書を題材とした技術者教育, 日本工学教育協会 平成 23 年度工学教育研究講演会講演論文集, 2011
- [13] 竹内広宜, 荻野紫穂, 中田武男, 坂本佳史, 福岡直明: テキスト分析技術を用いた開発関連文書の品質分析, 組込みシステムシンポジウム 2009, pp 93-100 , 2009
- [14] Arthur, J. D. and Stevens, K. T., Document quality indicators: A framework for assessing documentation adequacy. Journal of Software Maintenance: Research and Practice, 4: 129-142. 1992
- [15] 藤田 悠, 山本 雅基, 中澤 達夫, 塩谷 敦子, 池田貴一, 楡井雅巳: 組込みソフトウェア開発文書診断法, 情報処理学会研究報告, Vol.2010-EMB-16, No.37 , 2010
- [16] 森田良行: 基礎日本語 2—意味と使い方, 角川書店, 1981