

reqchecker:IEEE 830の品質特性に基づく日本語要求仕様書の問題点検出ツール

林 晋平 有賀 顕 佐伯 元司

電子情報通信学会論文誌 2018 p57 ~ p67

目次

- 背景,目的
- 提案手法,ツール
- RQ
- 分析方法
- 実験結果
- まとめ,所感

背景

- 要求仕様書は主に自然言語で記述されているため,文意のあいまい性などの問題がある.
- 品質の高い要求仕様書を得るためには,要求分析者がこの問題点を認識し,発見することが重要である.

目的

- **IEEE 830** で定義された品質特性をもとに,日本語で記述された要求仕様書の問題点を検出する手法及びその自動化ツール「reqchecker」を提案する.

提案手法

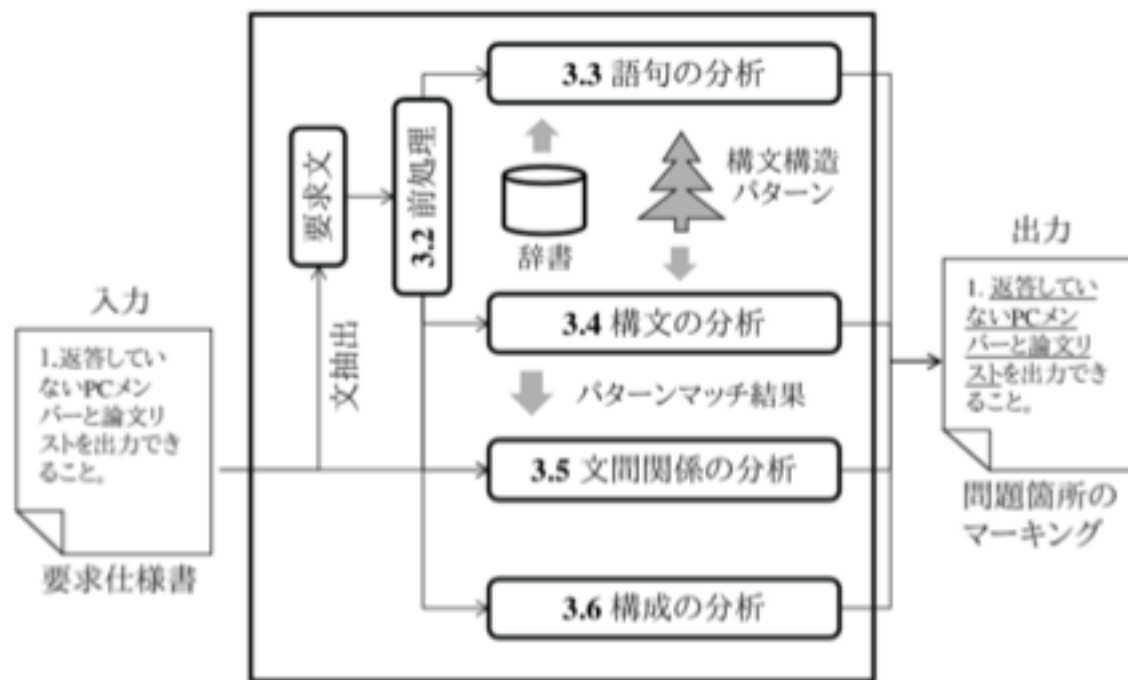


図 1 提案手法の概要

Fig. 1 Overview of our technique.

提案手法

- 語句の分析

得られた要件文を形態素に分解する.



語句の表記ゆれの問題を解消するため,既存の類義語の辞書を用いて,類義語の集合に変換する.

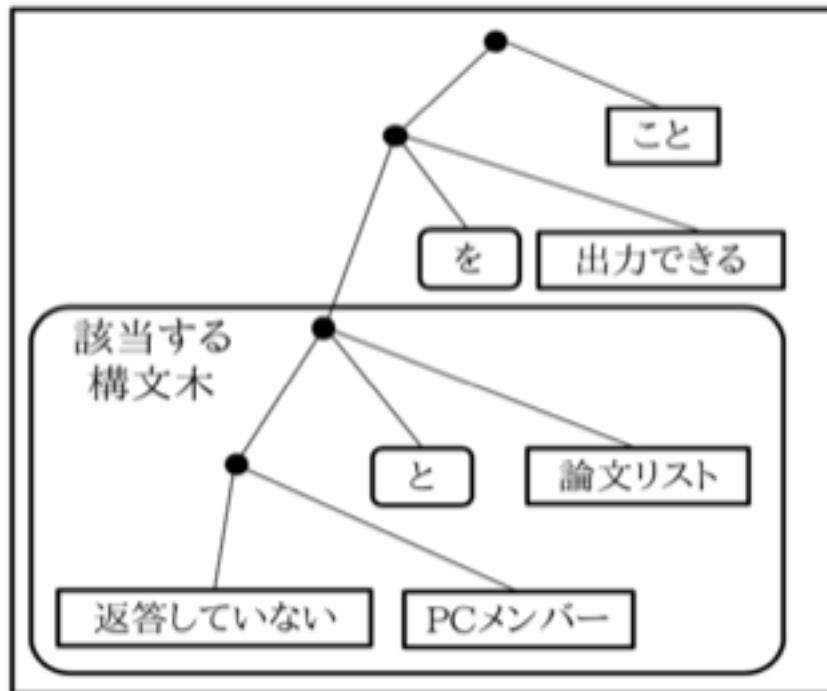


得られた類義語の集合と,辞書として用意しておいた品質特性の問題となりうる語句を比較する.

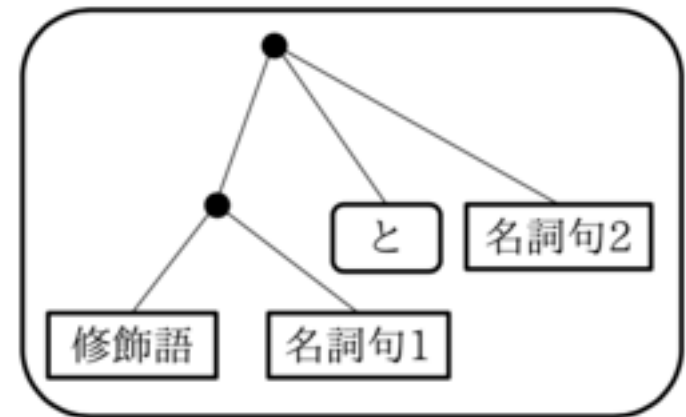
提案手法

- 構文の分析

(例文)返答していないPCメンバーと論文リストを
出力できること。



(a) 構文解析結果



(b) 複数解釈パターン 1

図 2 パターンマッチの例

Fig. 2 Example of the pattern matching.

提案手法

- 文間関係の分析

要求仕様書内の文と文を比較し,問題となりうる関係を検出する.

(例)無矛盾性

「熟成開始から24ヶ月までに出荷されるものは,
クリアンサと呼ばれる。」

「30ヶ月目に出荷されるものは,クリアンサと呼ばれる。」

提案手法

- 構成の分析

要求仕様書中の要求文,節,章の構成を把握して,品質特性の問題点を検出する.

(例)完全性[用語の定義がされていない]

要件仕様書内で,「定義」などの特定の名称のついた見出しが存在するかを調べる.

IEEE 830の品質特性

表 1 検出対象の品質特性項目と利用する検出手法

Table 1 Problem types of each quality attribute and how they are detected.

品質特性	項目	品質低下要因	検出手法
非あいまい性	a	語句の係り受け関係が複数ある	構文の分析
	b-1	主語が省略されている	構文の分析
	b-2	目的語が省略されている	構文の分析
	c	指示語が含まれている	語句の分析
	d	判断基準があいまいな語句を含む	語句の分析
	e	範囲や境界を表す語を含む	語句の分析
完全性	c-2	用語の定義が記載されている	構成の分析
無矛盾性	b	定義が矛盾している	構文の分析, 文間関係の分析
検証可能性	a	定性的な表現を含む	語句の分析
変更可能性	a-2	索引が付けられていない	構成の分析
	b	同じ要求が 2 箇所以上に現れる	文間関係の分析
	c	要求が互いに依存している	文間関係の分析
追跡可能性	a	章節番号が付与されていない	構成の分析
	c	各要求にユニークな番号が振られていない	構成の分析

問題点検出ツール(reqchecker)

「運賃を計算する際には、到着日と同じ日に乗り継ぐ場合は、出発地から目的地までの総飛行マイル数で運賃を算出する。」

要求文に識別番号がありません。
追跡可能性cに問題があります。

主語が省略されている可能性があります。
非あいまい性bに問題がある可能性があります。
「運賃を計算する際には、到着日と同じ日に乗り継ぐ場合は、出発地から目的地までの総飛行マイル数で運賃を算出する。」

目的語が省略されている可能性があります。
非あいまい性dに問題がある可能性があります。
「運賃を計算する際には、到着日と同じ日に乗り継ぐ場合は、出発地から目的地までの総飛行マイル数で運賃を算出する。」

あいまいな言葉が含まれています。
非あいまい性eまたはfに問題がある可能性があります。
「運賃を計算する際には、到着日と同じ日に乗り継ぐ場合は、出発地から目的地までの総飛行マイル数で運賃を算出する。」

あいまいな言葉・定性的な表現が含まれています。
非あいまい性dまたはeと関連しています。
検証可能性aに問題がある可能性があります。

修正が必要である可能性があります。

原文

問題となっている理由

問題箇所

問題となっている
品質特性と項目

(a) 問題点の一覧

要求文に識別番号がありません。
追跡可能性cに問題があります。

ある要求にユニークな番号が振られているかどうか

あいまいな言葉が含まれています。
非あいまい性dまたはeに問題がある可能性があります。

(b) 品質特性項目の説明

RQ

- RQ1

「reqchecker」の問題点検出精度はどのくらいか.

- RQ2

要求分析者は「reqchecker」を便利と感じるか.

RQ1の分析方法

- RQ1の評価実験には以下の例題を用いることとする.
- 例題には,意図的に品質特性に関わる問題点が含まれていることとする.

表 3 評価に用いた例題

Table 3 Examples used for the evaluation.

記号	名称	要求文数
R_{winery}	ワイン予約システム	14
R_{airline}	航空チケット予約システム	15
R_{meeting}	会議室予約システム	14

RQ1の分析方法

- 例題として用意した要求仕様書をツールに入力し,ツールから得られた指摘を正解と比較して適合率と再現率を算出する.
- 適合率Pと再現率Rは要求仕様書中に存在する問題点の集合 C とツールの出力した問題点の集合 D を用いて以下のように求める.

$$P = \frac{|C \cap D|}{|D|}, \quad R = \frac{|C \cap D|}{|C|}$$

RQ1の実験結果

表 4 適合率と再現率
Table 4 Precision and recall values.

品質特性	項目	R_{winery}				$R_{airline}$				$R_{meeting}$				全体			
		$ C $	$ D $	P	R	$ C $	$ D $	P	R	$ C $	$ D $	P	R	$ C $	$ D $	P	R
非あいまい性	a	2	1	1.00	0.50	2	0	NA	0.00	1	0	NA	0.00	5	1	1.00	0.20
	b-1	6	5	0.80	0.67	6	1	1.00	0.17	7	5	1.00	0.71	19	11	0.91	0.53
	b-2	12	21	0.52	0.92	3	15	0.20	1.00	4	11	0.27	0.75	19	47	0.36	0.89
	c	5	5	1.00	1.00	3	1	1.00	0.33	4	4	1.00	1.00	12	10	1.00	0.83
	d	5	7	0.57	0.80	17	21	0.33	0.41	12	9	0.67	0.50	34	37	0.46	0.50
	e	2	2	1.00	1.00	2	1	0.00	0.00	5	2	1.00	0.40	9	5	0.80	0.44
完全性	c-2	1	1	1.00	1.00	1	1	1.00	1.00	1	1	1.00	1.00	3	3	1.00	1.00
無矛盾性	b	1	0	NA	0.00	0	0	NA	NA	0	0	NA	NA	1	0	NA	0.00
検証可能性	a	1	8	0.13	1.00	4	11	0.36	1.00	3	9	0.22	0.67	8	28	0.25	0.88
変更可能性	a-2	1	1	1.00	1.00	1	1	1.00	1.00	1	1	1.00	1.00	3	3	1.00	1.00
	b	1	0	NA	0.00	1	0	NA	0.00	0	1	0.00	NA	2	1	0.00	0.00
	c	3	3	0.67	0.67	2	1	1.00	0.50	3	3	0.67	0.67	8	7	0.71	0.63
追跡可能性	a	1	1	1.00	1.00	0	0	NA	NA	0	0	NA	NA	1	1	1.00	1.00
	c	14	14	1.00	1.00	14	14	1.00	1.00	6	6	1.00	1.00	34	34	1.00	1.00
合計		55	69	0.68	0.85	56	67	0.49	0.59	47	52	0.62	0.68	158	188	0.60	0.71

$$P = \frac{|C \cap D|}{|D|}, \quad R = \frac{|C \cap D|}{|C|}$$

※NAは測定不能

RQ1の実験結果

- NAであったものを除いて,適合率は13項目中8項目
再現率は14項目中7項目で0.8以上となった.
- 適合率,再現率が0.8を上回った15項目中,6項目が
「非あいまい性」,4項目が「追跡可能性」に関する
ものであったことから「reqchecker」はこの2つ
の品質特性について高精度の検出能力を持っている.

RQ2の分析方法

- 情報工学を専門とした学生3名を被験者とする.
- RQ1で使用した要求仕様書を2つ配布し,ツールを使用した場合としなかった場合とで,品質特性に関わる問題点を発見,修正させる.
- 作業終了後被験者3名にアンケート調査を行う.

RQ2の実験結果

各項目において1から5の5段階評価を行う.

表 6 アンケート回答結果
Table 6 Result of questionnaire.

質問項目	回答	平均値
reqchecker が便利だったか	3, 4, 4	3.67
非あいまい性について役に立ったか	5, 5, 5	5
完全性について役に立ったか	4, 5, 5	4.67
無矛盾性について役に立ったか	2, 3, 4	3
検証可能性について役に立ったか	3, 4, 4	3.67
変更可能性について役に立ったか	2, 3, 5	3.33
追跡可能性について役に立ったか	3, 4, 4	3.67

RQ2の実験結果

- 全項目において評価の平均値は3を超えており、「reqchecker」についても個々の品質特性についても好評価であった。

まとめ

- 今回は予約システムに絞って実験を行ったため,他の領域の仕様書でも適応できるかは不明.
- 評価実験の被験者が3名のみであったため,定量的な評価が行えたとは言い難い(被験者を増やす必要がある).
- 一部の品質特性の問題点の検出に関しては,概ね良好であった.

所感

- 仕様書内の一文だけを見て分類するのではなく、他の文とも比較して分析をしていたことに驚いた.
- 形態素解析を用いた分類方法は勉強になった.
- 予約システム以外での結果も気になる.