

# 要求仕様のあいまいさを解消するための プロセスフレームワークの提案

中島 千壽<sup>†</sup> 高橋 宏季<sup>†</sup> 位野木 万里<sup>†</sup>

工学院大学<sup>†</sup>

## 1. はじめに

著者らの研究グループでは、自然言語で記述された要求仕様書の高品質化の研究に取り組み[1]、その一環として、あいまい表現となる同義語を自動的に指摘するツールの開発を試みてきた[2]。その過程で、要求仕様書を分析し、同義語辞書を自動で生成する手法を開発した。著者らは従来、設計要素のうち振る舞い用語に着目してきたが、アクター、データ、画面の同義語が、要求仕様をあいまいにする場合もある。

本研究では、他の設計要素の同義語にも適用可能とするために振る舞い用語の同義語の抽出過程を一般化し、様々な設計要素の同義語抽出に対応した要求仕様のあいまいさを解消するプロセスフレームワークを提案する。

以下、本稿は次のように構成する。2章では前提知識となる同義語および同義語抽出の考え方を説明する。3章では要求仕様のあいまいさを解消するためのプロセスフレームワークを提案する。4章では本プロセスフレームワークの利用と改善について考察し、5章で本稿をまとめる。

## 2. 要求仕様書における同義語と同義語表現

### 2.1 同義語について

同義語とは、意味は同じだが表現が異なる用語である。実際の要求仕様書で使用されていた同義語の例にすると、仕様書内では「格納」「保管」が同一の意味で使われていた。しかし、格納と保管は類義語ではない。このように、要求仕様書には特有の同義語が存在し、このような同義語を見つけるにはどのような用語が同義語に相当するか分析する必要がある。

### 2.2 Word2vec を用いた同義語辞書自動生成手法

著者らの研究グループでは、自然言語で記述された要求仕様書の高品質化のため、単位ベクトル表現化ツールである word2vec のサポートによって同義語辞書の作成作業をサポートする同義語辞書自動生成手法を開発した[3]。この手法では単語ベクトル間のコサイン類似度を計測、閾値以上の値を持つ単語ベクトルの要素に対応する単語の和集合を同義語として抽出する。

しかし、本手法はまだ洗練がされておらず同義語辞書の生成に多大なコストがかかっている。また、著者らは今まで設計要素のうち、振る舞い用語に着目してきたが、アクター、データ、画面の同義語が要求仕様をあいまいにする場合がある。

## 3. 要求仕様のあいまいさを解消するためのプロセスフレームワーク

仕様を構成する設計要素である、アクター、データ、画面も含めた同義語抽出の方法を検討し、要求仕様のあいまいさを解消するためのプロセスフレームワークを図1に

示す。図1は、要求仕様のあいまいさを解消するための検証や改善に必要な知識、観点、手続き、ツールの関係性を示したものである。

図1(a)は、要求仕様のあいまいさを解消するために必要と考えられる知識の概念を定義した。当該知識は、要求仕様のあいまいさを解消するプロセスから構成され、要求仕様の構成要素や特性に応じて、学習、指摘、改善の手続きが必要であることを示す。

図1(b)は、図1(a)で定義したあいまいさを解消するための知識の観点のバリエーションを示す。例えば、あいまいさに関する知識であれば、同義語、派生語、指示語、未定義語、主語のない動詞などが、あいまいさを誘発するリスクとなることを示している。また、ここでは、あいまい表現の学習には、組織内で知識を共有すること、機械学習の仕組みを利用することなどが考えられることも示している。さらに、あいまいな表現の改善に関しても、組織内で実施するレビューの他、各種ツールや、著者らが開発してきた要求仕様の一貫性検証支援ツールの活用などの対応のバリエーションがあることも示している。

図1(c)は、図1(b)の知識やプロセスのバリエーションに関する具体例を示している。図1(c)の左側は、振る舞い用語の同義語の概念の形式知化の事例であり、図1(c)の右側は、各種ツールを用いた同義語の特定方法の形式知化の具体例である同義語抽出のフローを示している。

このように、あいまいさを解消するための方法は多数ある。図1のプロセスフレームワークを活用することで、様々な方法の比較や使い分けの指針を整理しやすくする。

## 4. プロセスフレームワークの利用と改善に向けて

著者らは、図1(c)で示した同義語抽出の考え方と手法を2つの文書に適用した。適用した文書は、エンタープライズ系システムの各要求仕様書[4][5][6]のシナリオ記述を分析し抽出した振る舞い用語と、livedoor ニュースコーパスのトピックニュース[7]の記事を分析して抽出した振る舞い用語である。

表1に、要求仕様書から作成した同義語辞書の結果の抜粋を示す。表1では、No.1の「アップ」「格納」「保管」「保存」、No.2の「リリース」「改善」「改訂」No.4の「整理」「分類」といった一般の類義語辞書にはないものを指摘することができた。

表2にニュース記事から作成した同義語辞書の結果の抜粋を示す。表2では、No.1の「反応」「対応」、No.3の「コメント」「指摘」といった一般の類義語辞典にはないものを指摘できた。

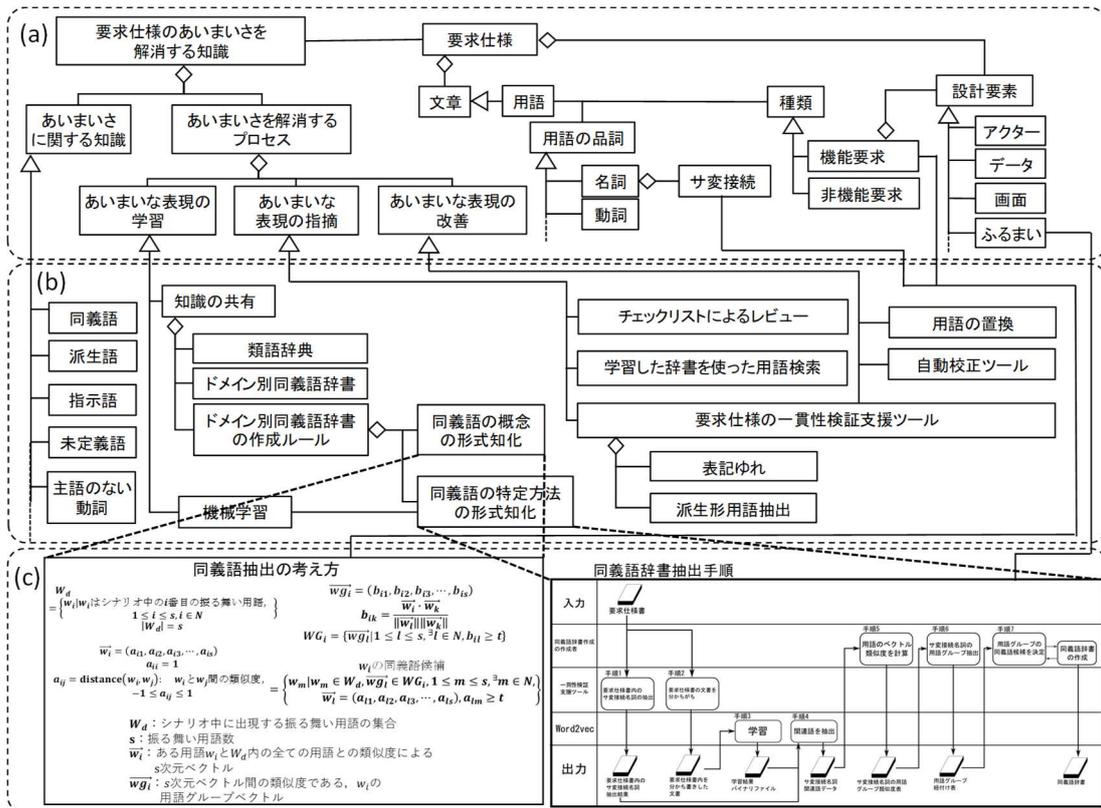


図1 プロセスフレームワークの構成

表1 要求仕様書による同義語辞書抽出結果

No.	同義語
1	アップ けんば 開始 格納 監査 受渡し 設置 徴収 保管 保存
2	イメージ リリース 改善 改訂 提言 保持 要望
3	メール 稼働 検証 交付 指定 支援 実施 収録 周知 進達 進捗 送検 送付 蓄積 提出 把握 配布 表示 返戻 郵送 履行
4	一括 継続 失権 整理 適用 認証 分類
5	改修 該当 修正 追加
6	開発 減少 設計 存在 低下

表2 ニュース記事による同義語辞書抽出結果

No.	同義語
1	イメージ 疲労 意見 意味 行動 支持 存在 対応 反応 理解
2	インタビュー 取材 代表
3	コメント バッシング 殺到 反論 否定 批判 評価 指摘 活動 期待 懸念 抗議 主張 心配 続出 展開 非難 復帰 擁護
4	メール 告白 撮影 謝罪 宣言 発覚 発生 展開 判明 浮上 報告
5	議論 掲載 交際 公開 削除 紹介 注目 投稿 発表 暴露
6	決定 出演 登場 参加 言及 終了 騒動 調査 発売 報道

表1表2は、「ドメイン別同義語辞書」のインスタンスと位置づけられる。また、「ドメイン」という概念を図1(a)に設定すれば、表1表2のインスタンスの位置づけを、より明確化できると考えられる。

5. まとめ

本研究では、振る舞い用語の同義語の抽出過程を一般化し、様々な設計要素の同義語抽出に対応した要求仕様のあいまいさを解消するためのプロセスフレームワークを提案した。本プロセスフレームワークにより、どの設計要素やあいまいさの種類や、あいまいさの特性の形式化のパリエーションの位置づけを明確にすることが可能になった。

今後は、本プロセスフレームワークを実装して、要求仕様の高品質化に役立てていく。

謝辞

要求仕様の一貫性検証支援ツール開発に関わる研究は、独立行政法人情報処理推進機構技術本部ソフトウェア高信頼化センター (SEC : Software Reliability Enhancement Center)が実施した「2015 年度ソフトウェア工学分野の先導的研究支援事業」の支援を受けたものである。また、本研究は JSPS 科研費 JP16K00105, JP19K11907 の助成を受けた。

参考文献

- [1] 位野木 万里, 近藤 公久, 省略と修飾パターンを用いた用語不一致検証による要求仕様の一貫性検証支援ツールの実現と適用評価, 日本ソフトウェア科学会, コンピュータソフトウェア, Vol.35, No. 3, pp. 109-127, 2018.
- [2] 高橋 宏季, 野村 典文, 近藤 公久, 位野木 万里, 要求仕様書における派生系アクター自動抽出手法: 組織変更による影響対応への効果, 情報処理学会, ソフトウェアエンジニアリングシンポジウム 2018 論文集, pp. 121-129, 2018.
- [3] 伴 凌太, 高橋 宏季, 位野木 万里, word2vec を用いた同義語辞書自動作成手法の提案と適用評価, 情報処理学会, 第 81 回全国大会講演論文集, pp. 265-266, 2019.
- [4] 厚生労働省, 年金業務システム (個人番号管理サブシステム等 (2 次開発情報連携分)) に係る設計・開発等業務及びアプリケーションソフトウェア保守業務 調達仕様書(案) <http://www.mhlw.go.jp/sinsei/shotatu/shotatu/shiyousho-an/160428-1.html> (参照 2020-01-10)
- [5] 厚生労働省, 労働保険適用徴収システムに係るシステム運用業務一式 <http://www.mhlw.go.jp/sinsei/shotatu/shotatu/shiyousho-an/120329-2.html> (参照 2019-01-10)
- [6] 厚生労働省, 「労働基準システムに係るアプリケーションプログラム回収業務一式 (平成 28 年度)」 <http://www.mhlw.go.jp/sinsei/shotatu/shotatu/shiyousho-an/151204-1.html> (参照 2019-01-10)
- [7] livedoor ニュースコーパス <http://www.rondhuit.com/download.html#1dce> (参照 2020-01-10)