

## 要求定義支援のための要求適合度を用いた事例検索

藤田 幸之<sup>†</sup> 金子 伸幸<sup>†</sup> 中元 秀明<sup>†,††</sup> 小川義明<sup>†,†††</sup> 阿草 清滋<sup>†</sup>

<sup>†</sup>名古屋大学大学院情報科学研究科 <sup>††</sup>株式会社野村総合研究所  
<sup>†††</sup>株式会社中電シーティーアイ

本稿では、ビジネスオブジェクトを利用して仕様化された要求仕様に対して、現在の開発要求との類似度を評価する指標として要求適合度を定義する。ビジネスオブジェクトは、分析から実装まで一貫したモデルを利用してソフトウェアを開発するアプローチである。自然言語で記述された要求記述に対して、事例に含まれる特徴的なビジネスオブジェクトとビジネスオブジェクトの構造を基に、要求適合度を計算する。要求適合度に基づき事例を検索し、ビジネスオブジェクト単位で修正を行い要求仕様を定義することで、分析段階の再利用を支援できる。本手法の評価として、保険業務に対する要求適合度の計算結果に基づき、要求適合度の妥当性について評価する。

## Case Retrieval for Requirements Definition Using Requirement Relevance

Hideyuki Fujita<sup>†</sup> Nobuyuki Kaneko<sup>†</sup> Hideaki Nakamoto<sup>†,††</sup> Yoshiaki Ogawa<sup>†,†††</sup>  
Kiyoshi Agusa<sup>†</sup>

<sup>†</sup>Graduate School of Information Science, Nagoya University  
<sup>††</sup>Nomura Research Institute, Ltd. <sup>†††</sup>ChudenCTI Co.,Ltd.

In this paper, we define requirement relevance in the requirement specification which specified by using the business object as an measure that evaluates the degree of relatedness with a present development requirement. The business object is an approach for the development of software by using a consistent model from the analysis to implementation. We calculate requirement relevance for the requirement description described by the natural language based on the structure of a business object and the discriminative business object contained in the case. We can support reusing in the analysis phase by retrieving the case based on the requirement relevance, correcting with each business object, and defining the requirement specification. We evaluate the fairness of the requirement relevance as an evaluation of this method based on the calculation of the requirement relevance to the insurance service.

### 1 はじめに

システム開発において要求記述から要求仕様を作成する要求仕様化は重要な工程である。要求仕様の誤りは開発の手戻りを引き起こし、開発コストの増大に繋がる。しかし、一般に顧客の要求は非常に複雑であったり、曖昧であったり、矛盾している場合が多い [1] ため、誤りの無い要求仕様の作成は非常に困難である。特にビジネスシステムの要求には、業務特有の規則や暗黙の理解が多く潜んでいる。このため誤りの無い要求仕様の作成には、要求分析者の経験や洞察力が必要である。

ビジネスシステムの開発において再利用指向の開発手法は、開発のコストを削減し信頼性を向上するために有効である。しかし、再利用指向の開発を行うには再利用するソフトウェア部品の選択を適切に行わなければならないため、多くの研究がなされている。

ビジネスオブジェクトを用いた要求定義手法 [2] では、要求仕様の再利用を行うことで設計仕様、ソフトウェア部品の再利用を行う。再利用のために要求記述から過去の要求仕様である事例の検索を行うが、事例を評価する基準が無く、参考にした事例だけを検索できないという問題が存在する。

本稿では、この問題を解決するために要求適合度を定義し、事例の検索手法を提案する。

事例の検索手法は、以下の前提の下で要求記述を用いて要求に適合する事例を検索できる。

- 同じ業務概念は必ず同名で表される
- 一つの名前で複数の業務概念を表さない

### 2 事例を用いた要求定義手法

業務類似性に基づく独自のビジネスオブジェクトを用いたソフトウェア再利用開発手法として、DBOA(Dual Business Object Approach)[2][3] があ

る。DBOAは実世界の業務概念とコンポーネントの間のギャップを埋め、両者をうまく結びつけるために考案された。

DBOAではRBO(Real Business Object)と呼ばれる実世界上で業務の概念を表すオブジェクトと、RBOをコンポーネントの組み合わせで実現するためのVBO(Virtual Business Object)と呼ばれるオブジェクトを用いることで、ソフトウェア開発の要求仕様と設計仕様を得る。

### 2.1 RBO:要求定義のためのビジネスオブジェクト

要求は主に機能要求、非機能要求、開発手順に関する要求の3つに分類できる。[4]

2.2節で説明するRBOを用いた要求定義手法で扱うのは機能要求の仕様と非機能要求の仕様である。RBO図はRBOを組み合わせで記述する。RBO図を用いて記述する要求は機能要求である。

#### 2.1.1 RBOの種類

RBOは業務中の要素と対応して、業務機能を表すFO(Function Object)、データ管理機能を表すDMO(Data Management Object)、業務データを表すDO(Data Object)の三種類に分類できる。RBOは階層構造を持っており、RBOを表現するために他のRBOを利用することができる。

##### FO(業務機能)

業務を構成する比較的独立した機能を表すオブジェクトである。FOは、外部とやりとりするインタフェースとして一連の業務機能を実行するメソッドを一つ以上持っている。メソッドが実行する一連の業務機能は、それ以上分割すると意味がなくなるような元素機能(データの生成、修正、チェックのいずれか)または他のFOやDMOのメソッドの呼び出しから構成される。

##### DMO(データ管理機能)

業務で永続的に使われ、管理されるデータを表すオブジェクトである。DMOは、外部とやりとりするインタフェースとして一連のデータ管理を行う業務機能を実行するメソッドを一つ以上持っている。メソッドが実行する一連の業務機能は、元素機能(データの挿入、更新、検索、削除のいずれか)または他のFOやDMOのメソッドの呼び出しから構成される。FOとDMOはどちらも業務における機能であり、互いに呼び出しが可能である。

##### DO(業務データ)

業務で使われる一時的なデータを表すオブジェクトであり、FOやDMOによって処理される。DOは、内部情報としてそれ以上分割すると意味がなくなるような元素データ(文字列、数、日付、真偽値のいずれか)または他のDOといったデータ項目から構成される。

#### 2.1.2 事例(RBO図)

システム全体を表す。事例は、呼び出すことのできるメソッドと、システムが実際に行うプロセス、業務データの流れを持っている。アクタ(システムの利用者)や事例内の利用できるメソッドの間で行う

業務データのやりとりを連結と呼ぶ。プロセスは連結の系列である。

事例の例が、図1の保険申込システムである。保険申込システムは、保険見積プロセスと保険申込プロセス、顧客確認プロセスの3つのプロセスを行えるシステムで、4つのメソッドが利用できる。

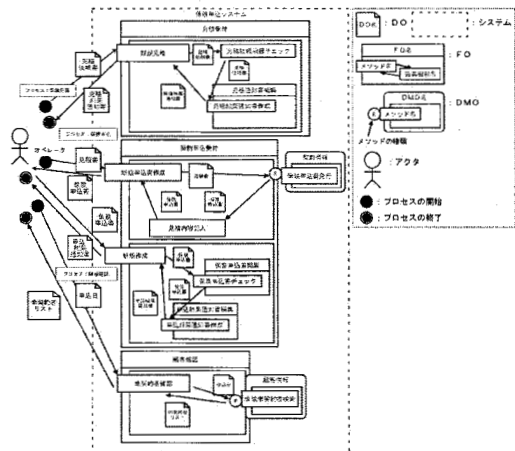


図1: 保険申込システムの事例

### 2.2 RBOを用いた要求定義手法

DBOAにおける要求定義フェーズにあたるRBOを用いた要求定義の流れを図2に示す。

RBOはリポジトリで管理し、要求定義において利用する。過去にRBOで作成された要求仕様を事例と呼ぶ。事例もRBO同様リポジトリで管理し、要求定義で利用する。

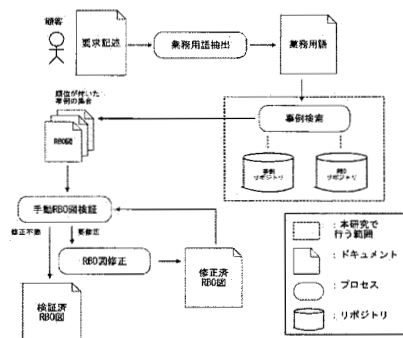


図2: RBOを用いた要求定義の流れ

RBOによる要求定義では、RBOリポジトリと事例リポジトリを利用して要求定義を行う。要求記述を基に要求獲得を行い要求を整理する。要求定義手法で扱う要求記述はそのドメインの用語を用いて文章で整理されたものであることが前提となる。

具体的には次のようなプロセスで要求定義を行う。

1. 業務用語抽出  
要求記述からドメインの用語として使われている業務用語を抜き出す。抜き出した業務用語は事例検索の検索語として利用する。
2. 事例検索

1. で抜き出した業務用語を検索語として事例の検索を行う。要求記述に対して参考にするべき事例の候補を得ることができる。事例はRBO図で書かれている。
3. RBO図検証
  2. で得られた事例のRBOの抜け、余分なRBOなどの要求との相違点について検証を行う。
4. RBO図修正
  3. の結果を基にRBO図を修正する。修正が終われば、再度RBO図の検証を行う。RBOの抜けや余りが無ければ修正は終了し、要求仕様は完成となる。

RBOによる要求定義では、ドメインの用語を用いて文章で整理された要求記述を与えると最終的にRBOで表された要求仕様が得られる。

### 2.3 RBOを用いた要求定義の問題点

RBOを用いた要求定義手法で行う事例検索において、事例を検索すると多くの事例が候補として得られてしまい、多くの事例を参考にしなければならないという問題があった。

本稿では、この問題を解決するために、要求記述を入力とし、要求に適合した事例を検索する手法を提案する。要求に適合した事例を検索する手法を実現するためには、クエリ(要求記述)と事例リポジトリが決まると一意に決まる指標を定義しなければならない。この事例が要求記述と適合しているかどうかの指標を本稿では要求適合度と呼ぶ。要求適合度の定義と検索手法については3章で説明する。

## 3 要求適合度による事例検索手法

2.3節で説明した問題点を解決するために、本手法では、要求記述から抽出した業務用語の集合をクエリとして、リポジトリに蓄積された事例の検索を行う。ある事例が要求記述で要求しているデータ、機能をどの程度満たしているかを表す指標として、要求適合度を定義する。この検索手法は以下のような性質を持つリポジトリに対して事例の検索ができる。

- 同じ業務概念は必ず同じ名前を持つ  
本検索手法は業務概念を名前で判断しているため、別名の同じ業務概念が存在すると検索結果に悪影響を及ぼす。
- 一つの名前で複数の業務概念を表さない  
上の性質と同様、名前から名前の表す業務概念が一意に決まらないと上手く検索出来ない。

### 3.1 要求適合度を定めるための要件

事例の検索を行うために、ある事例が要求記述で要求しているデータや機能をどの程度満たしているかを表すための指標を定義し、その指標を要求適合度と呼ぶ。要求適合度は、要求記述、事例の蓄積されているリポジトリの内容が決まると一意に決まる。

要求適合度を定めるにあたって、考慮すべき要件は以下の通りである。

- 希少度  
単語がどれだけ希少かを表す指標を定義する。本手法ではこの指標を希少度と呼ぶ。これは一般的な文書検索手法であるTF-IDF法[5]でIDFとして用いられている、希少なものほど重要であるという概念である。
- 業務概念の規模  
要求仕様を表す事例に含まれている実現すべき機能、データは一般に階層構造を持っている。要求仕様に現れる単語は、単語を表す概念の規模が異なるので、その違いを考慮しなければならない。規模が大きい概念ほど重要であると考えられる。
- 業務用語がリポジトリに含まれるかどうか  
要求記述の中には過去の事例で扱っていない新規業務を表す業務用語や、業務用語でない単語が含まれている可能性があるため、いずれかの事例に出現した単語のみを検索に利用する必要がある。

### 3.2 RBOの要求適合度

事例は業務機能を表すFOと、データ管理機能を表すDMO、業務データを表すDOの組み合わせで表現される。データに関する事例の要求適合度は、事例を構成しているDOの要求適合度を求めると決まる。機能に関する事例の要求適合度は、事例中でFOやDMOが実現する機能呼び出しているメソッドの要求適合度を全て求めると決まる。事例には機能を実行する順序や入力、出力するデータも書かれているが、それらを検索に反映するには要求記述を制限しなければならない。

DOの要求適合度は、DOを構成している要素である、利用しているDOの要求適合度と、元素データの要求適合度を求めると決まる。メソッドの要求適合度は、メソッドを構成している要素である、呼び出しているメソッドの要求適合度と、元素機能の要求適合度を求めると決まる。

要求適合度を求めるために、事例の特定の要素がどれだけ重要かを表す重要度という概念を定義する。要素Eの重要度 $I_E$ を以下の式で定義する。

$$I_E = f(R_E, S_E), R_E = g(C, 1/E_E) \quad (1)$$

ただし、 $R_E$ は要素Eを含む事例がどれだけ少ないかを表す希少度である。希少度 $R_E$ はリポジトリに蓄積された事例のうち要素Eを含む事例の割合が増加すると減少する。 $C$ はリポジトリに蓄積された事例の数、 $E_E$ は要素Eが含まれる事例の数を表す。 $g$ は $C, 1/E_E$ に関する増加関数である。

$S_E$ は要素Eがどれだけ多くの元素データ又は元素機能を持っているかを表す要素の規模である。 $f$ は $R_E, S_E$ に関する増加関数である。要素の重要度は希少度や規模が増加すると増加する。

本稿では、簡単のため $f(R_E, S_E) = R_E \times S_E$ 、 $g(C, 1/E_E) = C/E_E$ として実験を行う。

また、本稿ではRBOで表記された事例に関して以下の要件を考慮した。

- 機能とデータを別々に評価するため、システムが行う機能を表す FO、DMO とシステムで扱うデータを表す DO は別々に評価する。
- 自然言語の要求記述では、データがどの機能で利用されるかが明確に書かれていない可能性があるため、メソッドやメソッド内の元素機能に対する入力、出力 DO は検索に影響しないこととする。
- 事例に含まれる業務用語の差による正規化を行う。事例 A が事例 B を包含している場合、正規化を省くと事例 A は常に事例 B より要求に適合していることになる。必要のない要素が多い規模の大きな事例が検索結果の上位に現れてしまうことを防ぐために正規化を行う。

### 3.2.1 元素データの要求適合度

3.2 節で説明したように、データの重要度は、データの規模、希少度に関する増加関数である。元素データはこれ以上分割すると意味が無くなる最小単位のデータであるため、規模を 1 とする。元素データ a の重要度  $I_a$  を以下の式で定義する。

$$I_a = f(R_a, 1), R_a = g(C, 1/E_a) \quad (2)$$

ただし、 $R_a$  は元素データ a の希少度である。 $E_a$  は元素データ a が出現した事例の数、 $C$  はリポジトリに蓄積されている事例の数である。

元素データ a の要求適合度  $RR_a$  を以下のように定義する。

$$RR_a = \begin{cases} I_a & \text{クエリが元素データ名を含む} \\ 0 & \text{それ以外のとき} \end{cases} \quad (3)$$

### 3.2.2 データを表す RBO の要求適合度

2 章で説明したように、DO は元素データと他の DO を利用して構成されている。DO D の要求適合度  $RR_D$  を以下の式で定義する。

$$RR_D = \sum_{a_{dir} \in D} RR_{a_{dir}} + \sum_{d \in D} RR_d + RR_{D_{name}} \quad (4)$$

ただし、 $a_{dir}$  は DO D が直接利用する元素データ、 $d$  は利用している DO である。また、 $RR_{D_{name}}$  は DO の名前に関する要求適合度を表し、以下の式で定義する。

$$RR_{D_{name}} = \begin{cases} I_{D_{name}} & \text{クエリが DO 名を含む} \\ 0 & \text{それ以外のとき} \end{cases} \quad (5)$$

ただし、 $I_{D_{name}}$  は DO D の名前の重要度である。DO の名前の重要度は、その DO が存在すれば必ず利用している元素データの重要度の合計で決まり、以下の式で定義する。

$$I_{D_{name}} = \sum_{a_{dir} \in D} I_{a_{dir}} + \sum_{a_{ind} \in D} I_{a_{ind}} \quad (6)$$

$(a_{dir}, a_{ind} | R_{a_{dir}}, R_{a_{ind}} \leq R_{D_{name}})$

ただし、 $a_{ind}$  は DO が直接利用せず、利用している DO 内に存在する元素データである。DO D の名前の希少度  $R_{D_{name}}$  は、元素データ a の希少度 (式 2)

と同様にして求められる。DO の名前より希少度の高い元素データは特定の事例に特有の元素データであるため、DO が存在してもその元素データは含まれていないことが有り得るので、DO の名前の重要度を決定する際には用いない。

### 3.2.3 元素機能の要求適合度

データの場合と同様に、機能の重要度は、機能の規模、希少度に関する増加関数である。元素機能はこれ以上分割すると意味が無くなる最小単位の機能であるため、要素の規模を 1 とし、元素機能 a の重要度  $I_a$  を以下の式で定義する。

$$I_a = f(R_a, 1), R_a = g(C, 1/E_a) \quad (7)$$

ただし、 $R_a$  は元素機能 a の希少度である。 $E_a$  は元素機能 a が出現した事例の数、 $C$  はリポジトリに蓄積されている事例の数である。

元素機能 a の要求適合度  $RR_a$  を、以下のように定義する。

$$RR_a = \begin{cases} I_a & \text{クエリが元素機能名を含む} \\ 0 & \text{それ以外のとき} \end{cases} \quad (8)$$

### 3.2.4 機能を実現するメソッドの要求適合度

2 章で事例と FO、DMO について説明したが、事例から FO、DMO の持つ機能を利用するときには、FO や DMO そのものではなくメソッドを呼び出す。

また、メソッドから他の機能呼び出す場合も FO や DMO 全体ではなくメソッドを呼び出す。FO や DMO にメソッドがいくつあっても、一度に呼び出されるのは一つだけである。よって、事例の機能に関する要求適合度は FO 全体ではなく、メソッドの要求適合度を決めると決まる。

メソッド M の要求適合度  $RR_M$  を以下の式で定義する。

$$RR_M = \sum_{a_{dir} \in M} RR_{a_{dir}} + \sum_{m \in M} RR_m + RR_{M_{name}} \quad (9)$$

ただし、メソッド M が直接利用する元素機能を  $a_{dir}$ 、呼び出しているメソッドを  $m$  とする。また、 $RR_{M_{name}}$  はメソッドの名前に関する要求適合度を表し、以下の式で定義する。

$$RR_{M_{name}} = \begin{cases} I_{M_{name}} & \text{クエリがメソッド名を含む} \\ 0 & \text{それ以外のとき} \end{cases} \quad (10)$$

ただし、 $I_{M_{name}}$  はメソッド M の名前の重要度である。メソッドの名前の重要度は、そのメソッドが存在すれば必ず呼び出す元素機能の重要度の合計で決まり、以下の式で定義する。

$$I_{M_{name}} = \sum_{a_{dir} \in M} I_{a_{dir}} + \sum_{a_{ind} \in M} I_{a_{ind}} \quad (11)$$

$(a_{dir}, a_{ind} | R_{a_{dir}}, R_{a_{ind}} \leq R_{M_{name}})$

ただし、 $a_{ind}$  はメソッドが直接呼び出さず、呼び出している先のメソッド内に存在する元素機能のことを表す。メソッド M の名前の希少度  $R_{M_{name}}$  は、

元素機能 a の希少度 (式 7) と同様にして求められる。メソッドの名前より希少度の高い元素機能は特定の事例に特有の元素機能であるため、メソッドが存在してもその元素機能は含まれていないことが有り得る。よって、メソッドの名前の重要度を決定する際には用いない。

### 3.3 事例の要求適合度

DO からはデータに関する要求適合度を求め、メソッドからは機能に関する要求適合度を求める。ユーザが重視している項目に対して適合している事例を選ぶことができる。

事例検索を行う目的は参考にするべき事例を絞るためであるので、参考にするべき事例を 2 個程度に絞る。事例検索においては、検索手法における再現率や精度をそのまま利用することはできない。実際に参考にすべき事例が正しく決まることが最も重要である。

#### 3.3.1 データに関する要求適合度

事例内に含まれる全ての DO の要求適合度の合計が事例の要求適合度となる。ただし、DO の階層構造を考慮して他の DO の一部でしかない DO に関しては要求適合度を加算しないようにし、重複して加算を行わないようにする。

事例 C のデータに関する要求適合度  $RRD_C$  を以下の式で定義する。

$$RRD_C = \sum_{m_{dir} \in C} RR_{D_{dir}} / N_{Cd} \quad (12)$$

ただし、 $D_{dir}$  は事例内で用いられる DO (ただし、他の DO の一部でしかない DO は除く) である。また、 $N_{Cd}$  は正規化係数であり、以下の式で表される。

$$N_{Cd} = (1 - \text{slope}) \times \text{pivot}_d + \text{slope} \times l_{Cd} \quad (13)$$

slope は 1 次関数の傾きであり、 $\text{pivot}_d$  は事例に含まれるデータに関する平均の異なり単語数、 $l_{Cd}$  は事例 C に含まれるデータに関する異なり単語数である。

本稿では、文献 [5] に倣い slope  $\approx 0.2$  として実験を行った。データに関する要求適合度が最も高い事例がデータに関して要求を最も満たす事例となる。

#### 3.3.2 機能に関する要求適合度

アクタが直接利用するメソッドが決まると、事例が利用する機能が全て決まる。アクタが直接利用しているメソッドの要求適合度の合計が機能に関する事例の要求適合度となる。また、事例の規模に対して正規化を行うため、ピボット正規化 [5] の考え方をを用いて正規化を行う。

事例 C の機能に関する要求適合度  $RRF_C$  を以下の式で定義する。

$$RRF_C = \sum_{m_{dir} \in C} RR_{m_{dir}} / N_{Cf} \quad (14)$$

ただし、 $m_{dir}$  は事例 C が直接呼び出すメソッドである。また、 $N_{Cf}$  は正規化係数であり、以下の式で表される。

$$N_{Cf} = (1 - \text{slope}) \times \text{pivot}_f + \text{slope} \times l_{Cf} \quad (15)$$

slope は 1 次関数の傾きであり、 $\text{pivot}_f$  は事例に含まれる機能に関する平均の異なり単語数、 $l_{Cf}$  は事例 C に含まれる機能に関する異なり単語数である。

本稿では、データの場合と同様に slope  $\approx 0.2$  として実験を行った。

機能に関する要求適合度が最も高い事例が機能に関して要求を最も満たす事例となる。

### 3.3.3 要求適合度の統合

機能とデータのどちらを重視するべきか決まっていない場合に、要求適合度の統合を行う。事例に対して機能に関する要求適合度の順位と、データに関する要求適合度の順位が決まると、どちらかを重視しない場合の順位も決まる。具体的には、機能に関する要求適合度、データに関する要求適合度でそれぞれ順位付けし、最終的には Borda の手法 [6] を利用して順位を統合を行う。Borda の手法とは、複数の観点から評価を行ったとき、それらの評価を統合するための手法であり、各々の観点から順位付けを行い、順位合計が少ないものに高い評価を与える。

## 4 適用実験：保険業務に対する要求定義

3 章で提案した事例の検索手法に基づいて、保険業務中の契約締結業務に関するシステムの事例を検索する。

### 4.1 実験で利用した事例

事例を蓄積するリポジトリには申込方式 (コールセンター方式、代理店方式、インターネット申込方式) と保険の種類 (自動車保険、火災保険、生命保険) が異なる 9 個の事例が蓄積されている。事例に含まれる元素機能数の平均は 15.7 個、元素データ数の平均は 146.0 個である。なお、本評価実験で扱う保険業務は、実際の特定の保険会社のものではなく、保険業務に関する各種資料を基に作成したものである。

### 4.2 事例検索の例

クエリとなる要求記述を図 3 に示す。要求記述から抜き出した事例内で利用される業務用語を図 4 に示す。パラメータは重要度  $I_E$ 、希少度  $R_E$  を

$$I_E = f(R_E, S_E) = R_E \times S_E$$

$$R_E = g(C, 1/E_E) = C/E_E$$

slope を 0.2 とする。

以上の条件で事例検索を行った結果得られたそれぞれの事例に対しての要求適合度を表 1 に示す。コールセンター申込方式を C、代理店申込方式を A、インターネット申込方式を I と表す。機能を重視した場合はコールセンター方式の 3 つが最も要求と適合しており、データを重視した場合はコールセンター方式で自動車保険を扱う事例が最も要求と適合している。要求適合度の統合を行うと、コールセンター方式で自動車保険を扱う事例がデータに関しても機能に関しても 1 位なので、最も要求と適合している。

表 1: 事例検索の結果

申込方式	保険の種類	RRF <sub>C</sub>	順位	RRD <sub>C</sub>	順位	TF-IDF	順位
C	自動車保険	0.301	1	7.20	1	0.611	8
	火災保険	0.301	1	2.08	4	0.564	9
	生命保険	0.301	1	2.07	5	0.612	7
A	自動車保険	0.152	7	4.29	3	0.632	6
	火災保険	0.152	7	0.835	9	0.653	4
	生命保険	0.152	7	0.868	8	0.645	5
I	自動車保険	0.214	4	7.07	2	0.914	2
	火災保険	0.214	4	2.04	6	0.872	3
	生命保険	0.214	4	2.03	7	0.943	1

### 4.3 評価

本稿では、自動車保険を扱うコールセンター申込方式の事例が適合事例となる検索の実験を行う。

保険業務の保険申込業務の性質について述べる。保険申込業務の業務機能は、申込方式に依存して大きく変更される。また、保険対象が変わると見積方法が一部変わる。保険申込業務の業務データは、申込方式が変わると業務機能の増減に伴いデータが増減する。また、保険対象に依存して保険申込書、見積書など業務データの中身が大きく変わる。

検索で得られるべき適合事例はコールセンター申込方式の自動車保険を扱う事例である。リポジトリの内容から、最も望ましい適合度の順序は以下の順序となる。

- 業務機能に注目  
コールセンター申込方式の事例、見積業務がコールセンター申込方式と共通しているインターネット申込方式、代理店方式の順に適合事例と適合しているはずである。また、見積方法が保険対象によって変わるため、自動車保険を扱う事例は他の保険対象の事例に比べて適合事例と類似しているはずである。
- 業務データに注目  
自動車保険を扱う事例が最も要求と適合しているはずである。また、見積業務に依存する見積書などの業務データが存在するため、申込方式で比較するとコールセンター申込方式とインターネット申込方式の事例は代理店方式の事例に比べて適合事例と類似しているはずである。

#### 4.3.1 リポジトリの変化に伴う検索結果の変化

重要度  $I_E$ 、希少度  $R_E$  を

$$I_E = f(R_E, S_E) = R_E \times S_E$$

$$R_E = g(C, 1/E_E) = C/E_E$$

slope を 0.2 とおいて、自動車保険を扱うコールセンター申込方式の一般的な業務記述を入力として検索を行う。(検索語として用いた業務用語は図 4 の通り) この条件は 4.2 節で行った事例検索と同様である。

#### 初期条件のリポジトリ

結果は 4.2 節で行った事例検索と同様で、表 1 のように、コールセンター方式の自動車保険に関する

#### 業務概要:

A 社のホームページに保険の紹介を掲載する。  
顧客が、ホームページをみて保険に興味を持った場合、電話または、電子メールで A 社のコールセンターに保険資料請求を行う。  
コールセンターでは顧客の住所確認を行い、資料を送送する。  
顧客は資料を見て、コールセンターに電話したり、資料についている見積依頼書を送ったりして、保険の新規見積を依頼する。  
コールセンターのオペレータは、見積依頼内容をもとに見積書/保険申込書発行を行う。保険申込書には見積依頼の内容から計算した保険金や保険料が印刷される。コールセンターではその保険申込書と返信用封筒を顧客に郵送する。  
顧客は契約する場合は、保険申込書を使用して、保険料入金を行う。入金方法にはコンビニエンスストア、銀行、郵便局払い、クレジットカード払いがある。ただし、クレジットカード以外では分割払いはできない。(一回で全額支払う)  
入金情報は、コンビニエンスストア、銀行、郵便局払い、クレジットカード会社等から、A 社に送られる。その情報は、入金情報として管理される。  
顧客は保険申込書に署名、捺印のうえ、コールセンターに返信用封筒で送り返してくる。  
コールセンターでは、保険申込書チェックを行う。問題なければ、その内容で新規契約登録する。  
顧客が自分で保険申込書を訂正していることがあるので、その場合は、契約内容修正をする。また、入金情報を参照してすでに入金が行われているかをチェックする。未入金だと契約は成立しないので、顧客に連絡を取って確認する。  
コールセンターでは全てのチェック済の保険申込書をセットにして、保険会社の事務部門に送る。  
事務部門では、送られた書類を保管する。  
事務部門では、コンピュータセンタに連絡し、保険証券発行を依頼する。  
事務部門のプリンタにセットされた専用用紙に、前日の保険申込書に基づいて全ての保険証券が印刷される。保険証券交付の一覧表(対象証券番号と顧客名の一覧表)も印刷される。  
印刷された保険証券を契約者に郵送する。  
見積書/保険申込書を郵送して1ヶ月以上たっても申込みがない客には、コールセンターから電話をかけて様子を聞く。

#### 見積、申込に必要な項目の内訳:

申込日  
保険契約者 住所、氏名、生年月日、性別  
被保険自動車 車名、型式、排気量、車検満了日  
保険期間  
保険料  
担保種目 対人賠償、対物賠償、人身傷害、搭乗者傷害、盗難担保  
特記事項 運転者年齢条件、年間走行距離、免許証の色、用途・車種

図 3: 保険申込業務の要求記述

見積依頼書 新規見積 見積書 保険申込書発行 保険申込書  
保険料 新規契約登録 保険申込書チェック 保険証券発行 申込日  
保険契約者 住所 生年月日 性別 車名 型式 排気量 車検満了日  
保険期間 保険料 担保種目 対人賠償 対物賠償 人身傷害  
搭乗者傷害 盗難担保 特記事項 運転者年齢条件 年間走行距離  
免許証の色 用途・車種

図 4: 要求記述から抽出した業務用語

表 2: 目的事例が多く含まれるリポジトリに対する事例検索の結果

申込方式	保険の種類	RRF <sub>C</sub>	順位	RRD <sub>C</sub>	順位
C	自動車保険	0.267	1	4.20	1
	火災保険	0.267	1	1.71	6
	生命保険	0.267	1	1.78	4
A	自動車保険	0.143	7	2.38	3
	火災保険	0.143	7	0.746	9
	生命保険	0.143	7	0.795	8
I	自動車保険	0.191	4	4.13	2
	火災保険	0.191	4	1.68	7
	生命保険	0.191	4	1.75	5

保険申込システムを求めている要求記述の意図と一致する結果が得られた。

提案手法で事例検索を行った結果、4.2節で設定したリポジトリとクエリ、重要度、希少度、slope に対しては機能、データともにコールセンター申込方式の自動車保険を扱う事例が高い要求適合度を得て、参考にすべき事例が上位に現れることが確認できた。また、TF-IDF 法で検索した場合の順位は意図した並びと全く異なっていた。

#### 目的事例が多く含まれるリポジトリ

目的の事例と似た性質を持つコールセンター申込方式で自動車保険を取り扱う事例をリポジトリに 9 個追加し、重要度、希少度、slope に関しては同じ条件で検索を行った。追加した事例は FO「見積書編集」のメソッド「見積書作成」の元素機能「等級見積」を他の名前に変更したものである。リポジトリの変更により、目的の事例に含まれる要素の希少度は低下する。

変更したりリポジトリに対して提案手法で検索を行った結果を表 2 に示す。目的の事例が増えたことで、希少度の影響が小さくなった。データに関する要求適合度の順位は一部変わったが、これらは参考にすべき事例ではないので順位が変わっても問題ない。

#### 事例数が増加したりリポジトリ

リポジトリ内に蓄積されている事例 9 種類に対してそれぞれ元素機能を 1 つ変更した事例を 9 個追加し、リポジトリの規模を当初の 10 倍にし、パラメータは同じ条件で検索を行った。リポジトリの変更により、ある要素が事例に同じ回数だけ出現したとき、要素の希少度は 10 倍となる。

規模を 10 倍にしたリポジトリに対して 4.2 節の事例検索で最も事例と適合しなかった代理店方式の火災保険を取り扱う事例 (表 3 中では事例 (a) と呼ぶ) だけに新たに元素機能と元素データを 1 個ずつ加え、加えた元素機能と元素データを表す単語をクエリに含めて影響を調べた結果と、4.2 節で用いたりリポジトリに対して同様の操作を行った影響を調べた結果を表 3 に示す。

実験で使用したりリポジトリの元素機能数の平均は 15.7 個、事例に含まれる元素データ数の平均は 146.0

表 3: 元素機能、元素データの追加が検索結果に与える影響

対象事例	リポジトリ事例数	RRF <sub>C</sub>	RRD <sub>C</sub>
目的の事例		0.301	7.20
事例 (a)		0.152	0.835
(a) に要素追加	90	1.85	2.33
	9	0.321	0.980

表 4: クエリに「等級見積」を加えた事例検索の結果

申込方式	保険の種類	RRF <sub>C</sub>	順位	RRD <sub>C</sub>	順位
C	自動車保険	0.383	1	7.20	1
	火災保険	0.300	2	2.08	4
	生命保険	0.300	2	2.07	5
A	自動車保険	0.152	7	4.29	3
	火災保険	0.152	7	0.835	9
	生命保険	0.152	7	0.868	8
I	自動車保険	0.290	4	7.07	2
	火災保険	0.214	5	2.04	6
	生命保険	0.214	5	2.03	7

個で、9 倍強の差がある。この差とクエリにデータを表す業務用語が多く含まれていたことが検索結果に与える影響に現れており、元々規模が小さく、クエリに含まれる業務用語が少なかった機能についての要求適合度に関しては既存のクエリが無意味になってしまうほどの影響を与える。

表 3 の結果から、重要度を求める関数  $f$ 、希少度を求める関数  $g$  は事例の規模と事例数の割合によって希少度と事例の規模の影響のバランスを取るために関数を変更する必要があることが分かった。

#### 4.3.2 クエリの変化に伴う検索結果の変化

事例を特定できるクエリを作成したとき、求める事例が得られるかどうかを評価する。

##### 事例を特定する要素の不足

4.2 節で行った検索のクエリには自動車保険の業務機能を特定できる業務用語が不足していた。そこで、自動車保険を扱う事例特有の元素機能である「等級見積」をクエリに加えて検索を行う。検索の結果を表 4 に示す。コールセンター方式、インターネット申込方式の自動車保険を扱う事例の機能に関する要求適合度が上昇した。クエリから抜けていた自動車保険を扱う事例を特定するための業務用語を補うことで、業務機能に関して最も要求と適合した事例が特定できた。

## 5 関連研究

### 5.1 構造化文書の検索手法

#### キーワードを利用した検索手法

絹谷ら [7] は、キーワードをクエリとして XML 文書を検索する XRANK, XSearch, MPD などの手法を紹介している。このうち、XSearch では  $tf \cdot idf$  の変形の  $tf \cdot ilf$  を利用している。 $k$  をキーワード、 $n_k$  を葉ノード、 $occ(k, n_k)$  を  $n_k$  中に出現する  $k$  の頻度

とし、 $N$  をコーパス中の全ての葉ノード集合とすると、 $tf(k, n_i)$  と  $ilf(k)$  は、

$$tf(k, n_i) = \frac{occ(k, n_i)}{\max\{occ(k', n_i) | k' \in words(n_i)\}}$$

$$ilf(k) = \log\left(1 + \frac{|N|}{\{n' \in N | k \in words(n')\}}\right)$$

と表される。特徴としては、検索キーワードの出現位置からノード間の関連性を調べる点と、解答候補の文書の大きさをノード数で計測している点である。本稿で定義している要求適合度は RBO 図で書かれた事例に特化した指標であり、データと機能に分けて要求適合度を決定し、DO やメソッドの名前の意味を考慮した重みづけを行うなどの違いがある。

また、藤本ら [8] は、TF-IPF と呼ばれる指標を利用して単語の重みづけを行っている。TF は TF-IDF でも利用されている単語の出現頻度であり、IPF (Inverse Path Frequency) は XML 文書構造を考慮に入れた単語の特定性を表す尺度である。単語  $t$  の IPF は以下の式で計算される。

$$ipf_p(t) = \log \frac{N_p}{df_p(t)} + 1$$

本稿で対象としている RBO 図で書かれた事例に対しては、特定の部分文書内における単語の特定性は意味を持たないため、IPF は事例検索には適していない。

## 5.2 要求の検索

### ユースケースの用語を用いた相似構造計算による手法

Software Requirements Retrieval Using Use Case Terms and Structure Similarity Computation [9] では、ユースケース記述の検索手法を提案し、手法を実現するツールの実験を行っている。

この研究ではユースケース記述の 9 項目をクエリとしてベクトル類似度計算で検索を行っている。重みづけに TF-IDF 法の考え方を利用している点、ドキュメントの文字列を検索している点では本研究と類似しているが、各項目ごとに自然言語と自然言語を比較しており適合率が低くなる可能性が高い。本研究ではどの事例にも出現しない単語は検索クエリとして利用しておらず、不要な単語が原因となる検索結果の誤差の発生を回避することができる。

## 6 おわりに

本稿では、要求記述と過去の要求仕様である事例との要求適合度を定義することで事例の検索を行う手法を提案した。

過去の要求仕様である事例と要求記述との事例の階層構造を考慮した要求適合度を定義した。提案手法では、定義した要求適合度に基づき、要求記述に含まれる業務用語を入力として事例検索が行えることを示した。

保険業務の事例を蓄積したりポジトリを作成し、作成したりポジトリに対して適用実験を行い、提案手

法が要求記述と適合した事例を検索できることを示した。クエリに不足や誤りがある場合は求めている事例が検索できない可能性があることを確認した。

今後の課題を以下に示す。

### 検索手法の改善

RBO 同士の依存関係や共起関係などを利用した検索を行うことで、業務の流れを意識した検索を行えるようになる可能性がある。

### 要求記述のぶれへの対応

本稿の検索手法では事例内に出現する業務用語が要求記述に含まれていないと検索できない。シソーラスを利用することで、要求記述のぶれへ対応できるが、シソーラスを作成するコストが検索精度の向上に見合うかが課題となる。

### 謝辞

本研究を進めるにあたり、熱心に御指導、議論して頂きました阿草研究室、結縁研究室、小川研究室の皆様へ感謝致します。本研究の一部は文部科学省リーディングプロジェクト基盤ソフトウェアの統合開発 e-Society 高信頼 WebWare の生成技術および文部科学省科学技術研究費基盤研究 (B) 課題番号 17300006 の助成による。

### 参考文献

- [1] 大西淳, 郷健太郎, ソフトウェアテクノロジー 9 要求工学. 共立出版, 2002.
- [2] 赤川奈緒, 金子伸幸, 中元秀明, 小川義明, 阿草清滋. ビジネスオブジェクトリポジトリを用いた要求定義手法, 日本ソフトウェア科学会第 20 回大会論文集, p.1E-3, August 2003.
- [3] 金子伸幸, 赤川奈緒, 中元秀明, 小川義明, 阿草清滋. ビジネスオブジェクトを用いた要求分析・設計手法. 豊坂恒夫, 満田成紀 (編), ソフトウェア工学の基礎 X 日本ソフトウェア科学会 FOSE 2003. 近代科学社, 2003.
- [4] Merlin Dorfman, Richard H. Thayer (編). *Standards, Guidelines, and Examples on System and Software Requirements Engineering*. IEEE Computer Society Press, 1990.
- [5] 北研二, 津田和彦, 獅水堀正幹, 情報検索アルゴリズム, 共立出版, 2002.
- [6] 佐伯胖, 「きめ方」の論理: 社会的決定理論への招待, 東京大学出版会, 1980.
- [7] 絹谷 弘子, 波多野 賢治, 吉川 正俊, 植村 俊亮: キーワードを利用した XML 文書検索, 情報処理学会論文誌: データベース, Vol.45, No.SIG.7(TOD.22), pp. 255-273, 2004.
- [8] 藤本 圭, 清水 敏之, 波多野 賢治, 鈴木 優, 天竺 俊之, 絹谷 弘子, 吉川 正俊: “XML 情報検索システムとその高速化に関する研究”, 電子情報通信学会 第 17 回データ工学ワークショップ (DEWS2006) 論文集, 6B-i10, 2006 年 03 月.
- [9] Akadej Udomchaiporn, Nakornthip Prompoon, Pizzanu Kanongchaiyos, *Software Requirements Retrieval Using Use Case Terms and Structure Similarity Computation*, The 13th Asia Pacific Software Engineering Conference (APSEC06), Bangalore, India, Dec. 2006, pp.113-120.