

自然言語処理を用いた仕様書から VDM++仕様への変換支援について

酒井 玲弥
 仙台高等専門学校
sakainit@gmail.com

岡本 圭史
 仙台高等専門学校
okamoto@sendai-nct.ac.jp

要旨

ソフトウェアの不具合がもたらす悪影響は大きく、信頼性を高めるための設計手法の一つとして形式手法が用いられている。VDMは形式手法の一種であり、形式仕様記述言語 VDM++を持つ。VDMは実際の開発現場においても活用され、成果を挙げている。しかし、自然言語の要求をモデル化するプロセスにおいては、計算機による支援が進んでいない。そのため、人手による作業がほとんどであり、形式仕様の作成者への負担が大きいという問題がある。本論文では、形式手法導入コスト低減を目指した、自然言語で記述された要求仕様書から VDM++仕様への変換手法の提案を試みる。また、型定義、関数・操作定義に関する支援手法を実装し、評価を行う。評価の結果、人間の感覚に近い用語抽出を目指した、複合名詞の結合手法の有効性を確認することができた。しかし、定義する用語を効率的に探索することを目的とした抽出した語の並べ替え手法は、並べ替えを行わない場合と比較して精度は向上したが、改善の余地が残る結果となった。以上の結果をもとに VDM++で定義される用語の特徴について考察を行い、得られた知見を共有する。

1. はじめに

日常生活を支えるシステムがソフトウェア集約的になるにつれ、ソフトウェア技術への依存度が高まっている。また、近年のソフトウェアは規模の拡大、複雑化を続けている。そのため、ソフトウェアの不具合がもたらす悪影響は大きく、信頼性を高めるための設計手法が求められる。形式手法は、数理論理に基づいて設計された形式仕様言語による、曖昧さが排除された仕様記述と設計方法論、及び専門知識を総合し、システムに求められる信頼性の達成を目指すアプローチである。形式手法は、実際の開発現場においても活用されており、成果を挙げている[1]。VDM(Vienna Development Method)は、形式手法の一種であり、世界初の ISO 標準仕様記述言語である VDM-

SL と VDM++を持っている。これら VDM の形式仕様記述言語は VDM Tools[2]や Overture Tools[3]といったツールにより仕様の検証が可能であり、VDM Tools には C++と Java のコード生成機能がある。加えて、拡張機能として C#コードへの変換も開発されるなど、VDM++を開発現場に取り入れるための研究開発が活発に行われている[4][5]。また、ICカードの Felica ファームウェア開発プロジェクトでは、設計プロセスにおいて VDM++が適用されており、日本語での開発現場においても形式手法が適用されている[6]。以上のようにソフトウェアの開発工程に形式手法を適用することには、あいまいさの除去やツールによるテストの自動化など多くの利点があるが、形式手法を適用した開発の初期段階である、自然言語の要求をモデル化するプロセスにおいては、計算機による支援が進んでいない。そのため、人手による作業がほとんどであり、形式仕様の作成者への負担が大きいという問題がある。

以上の問題を解決するアプローチとして、VDM++仕様作成プロセスの一部自動化による変換支援手法を試みた。VDM++仕様作成を人手で行う際に発生する作業である、モデル化する用語の抽出作業を支援する。また、変換支援プロセス内で発生する抽出された単語のリストをもとに、見落としや自然言語の曖昧さに起因する曖昧な定義や冗長な定義を防止する。以上により、形式手法導入コストが低減され、プロジェクトへの形式手法導入が容易になるため、ひいては高品質なソフトウェア開発が可能になると考えられる。本論文の目的は、以上のアプローチに関する実験で得られた知見の共有である。モデル化する用語の抽出作業への支援として実装を行った、候補となる単語の抽出と並び替えの実験に関して、結果と考察を論じる。

2. 関連研究

2.1. VDM++仕様と自然言語仕様を併せた品質向上

大森らの研究では、仕様策定過程の全体を通じて、

自然言語で記述された仕様と形式手法による形式モデルの相互変換により、自然言語記述の品質向上を行う手法が提案された[8]。これにより、関係者の共通理解の得やすい自然言語記述と、厳密な形式モデルの長所を両立させることが可能となる。この手法では、自然言語と形式モデルの相互変換を頻繁に行うため、支援ツールの開発が行われた。支援ツールにおいては、キーワードのマーキングを行い、マークしたキーワードの定義を記述する用語辞書とキーワードの関連を構造化するデータ辞書の生成を補助する。そして、生成された辞書をもとに形式モデルのテンプレートが出力可能である。しかし、支援ツールにおけるキーワード抽出は手動で行う必要がある。手動によるキーワード抽出は、一般に使われる意味とは違う、対象ドメイン特有の意味が存在する単語に対応できるなどの利点があるが、大規模な要求仕様書に適用する場合、キーワードの抽出に大きな労力がかかると考えられる。そのため、本論文で提案する手法では、キーワードの候補となる単語、複合語を自動で抽出し、作業量を低減する。また、モデル化する単語を自動抽出された候補の中から人手により選択することで、手動による作業の利点を損なわない。

2.2. 機械学習を用いたアプローチ

執行らは、機械学習を用いた VDM++仕様生成へのアプローチとして、自然言語仕様から単語を抽出し、多項ロジスティック回帰分析を用いて型定義と定数定義への分類を行った[9]。また、菅らの研究では、執行らの手法に加え、日本語 WordNet を用いて、ある単語の下位概念との関係を表す概念レベルを算出し、クラスとインスタンス変数となる単語の分類を行った[10]。これらの手法は、関数・操作の定義に対応していないが、本論文では動詞の抽出と係り受け関係による支援手法を提案している。機械学習による単語の分類を行うにあたって、関連研究では、自然言語の仕様書を教師データとして、事前に学習済みモデルを用意していたが、仕様書の実態が変わると予測精度が低下するという問題がある。要求仕様書は公開されているものが少ないため、最適な教師データを手に入れられるとは限らない。本論文のアプローチでは、入力文書に対する自然言語処理で支援が完結するため、要求仕様書の分野や所持しているデータ量によらない支援が可能である。また、機械学習を用いた VDM++仕様の自動生成は、モデル作成者の作業量がかからず、知識が不要であるという利点があるが、モデル化に必要な単語を網羅するには至っておらず、誤った定義や定義漏れを避けられない。要求仕様の完全性という

観点において、自然言語仕様に記述されている仕様を網羅することは重要である。そのため、本論文では一部自動化による変換支援という異なるアプローチにより、人間の知識を一部介入させる。これにより、モデル化する単語の正確性が向上し、ひいては、モデル化する仕様の網羅性の向上が期待できる。

3. 変換支援手法のアプローチ

3.1. 変換支援手法の使用シナリオ

大森らの手法[8]では、VDM++を用いた開発フェーズとして、図 1 のような手順を想定している。

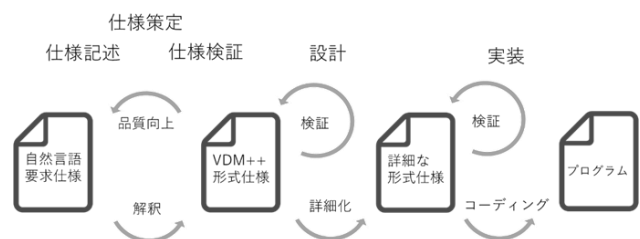


図 1 要求の洗練プロセス (大森 2010, p.20)[8]を参考に作成

本手法は、図 1 のプロセス中の自然言語仕様を VDM++へ変換する、仕様策定段階での使用を想定している。3.2 節以降で示す手法は、変換作業の一部自動化と必要な知識を作成者が補うことで、VDM++への変換を半自動的に行い、最終的に陰仕様の完成を目指すものである。陰仕様では、要求仕様書中に記述される「要求」と「仕様」をモデル化し、数理的な性質を満たしているか静的検証による仕様検証が可能である。

本論文においては、以上の一部実装として、型とシグニチャの抽出作業の支援に注力した。作成者は自然言語で記述された仕様書から、仕様をモデル化するために必要な単語を抽出する。抽出作業において、必要な単語の候補をシステムが自動で優先順位をつけて抽出し、作成者は型や関数になる単語を選択することで、モデリングが可能である。以上の手法を取るため、VDM++では、様々な抽象度での仕様記述が可能であるが、本手法によって作成される VDM++仕様は、自然言語仕様に含まれる単語を各要素に対応付けた、逐語訳的な変換をされたものである。よって本手法は、形式手法および VDM++の記述に関して最低限の知識があり、初歩的な VDM++仕様記述ができる技術者を対象としている。

3.2. 変換支援手法全体の流れ

変換支援手法全体の流れを図2に示す。入力された仕様書データを分析するため、形態素解析と構文解析を行う。そして、複合名詞の結合を自動で行い、類義語の統合を行うことで、後のモデル化に適した分かち書きデータを得られる。次に、分かち書きデータをもとに、型定義の支援を行う。型として定義される候補となる単語の抽出を自動で行い、型候補のリストをもとに人手により型の定義を行う。ここで、単語を無作為に提示すると人間が判断する上で効率が悪いので、型候補のリストは、型になる可能性が高い単語を上位に並べ替えたリストを提示する。ここまでで得られる、分かち書きデータと定義された型をもとに関数・操作定義の支援を行う。関数・操作の候補となる単語を自動で抽出し、得られた関数・操作候補のリストをもとに人手により関数・操作の定義を行う。関数・操作の定義も同様に最終判断を人手により行うため、候補のリストは、関数・操作になる可能性が高い単語を上位に並べ替えたリストを提示する。選択した型および関数・操作は辞書データとしてまとめられ、辞書データをもとにVDM++陰仕様が出来される。

3.3. 形態素解析・構文解析

はじめに、型定義や関数定義に必要な情報である、各単語の品詞情報を特定するため、形態素解析をツールの使用により行う。本論文では、形態素解析ツールとして、Sudachi を使用している。なお、Sudachi の辞書はデフォルトの Core、分割モードは最小単位で分割されるモードの A を選択している。構文解析は機械学習ベースの自然言語処理ツールである GiNZA[12]を使用した。GiNZA は spaCy と Sudachipy を用いて実装されており、spaCy の CNN マルチタスク学習器を用いて日本語

Universal Dependencies(以下 UD) の品詞ラベルと係り受けのラベルを付与する。また、GiNZA v5 では Transformers モデルにより、さらに高精度な構文解析を実現している。

3.4. 複合名詞の結合

要求仕様書はシステムの機能および、実現方法が記述された文書であり、一般的な文書と比較して専門性の高い用語や、システム独自の用語が用いられる可能性が高い。また、「話題沸騰ポット要求仕様書」[11]のVDM++仕様を人手により作成した際にモデル化した用語にも、複合名詞が多いという特徴が見られた。中川らは「専門用語の多くは複合語、とりわけ複合名詞であることが多い。」(中川 2003, p.27)[13]と述べており、専門性が高い文書である要求仕様書には複合名詞が多く含まれていると言える。以上より、要求仕様書をモデル化する際、複合名詞の用語を抽出することで人間の定義する型に近い用語抽出が行えるという仮説を立てた。また、専門用語の抽出に関して中川らの手法では、最小単位の名詞および複合名詞を候補として抽出し、優先順位付けを行うことで抽出を行っている[13]。VDM++において必要な単語は前述の通り、専門用語が多いと考えられるため、複合名詞の結合および、3.5 節で後述する並べ替えに関して、「話題沸騰ポット要求仕様書」に対し複数の実験を行い、型として定義される単語について考察する。

複合名詞の結合に関して、形態素解析の段階で、複合名詞を一つの単語として分割するアプローチと、最小単位に分割された単語を後から結合するアプローチが考えられる。形態素解析の段階で分割する場合は、各ドメイン知識に対応できる辞書をそれぞれ構成する必要がある、現状では困難である。そのため本論文では、形態素解析により最小単位に分割された単語をもとに、後から

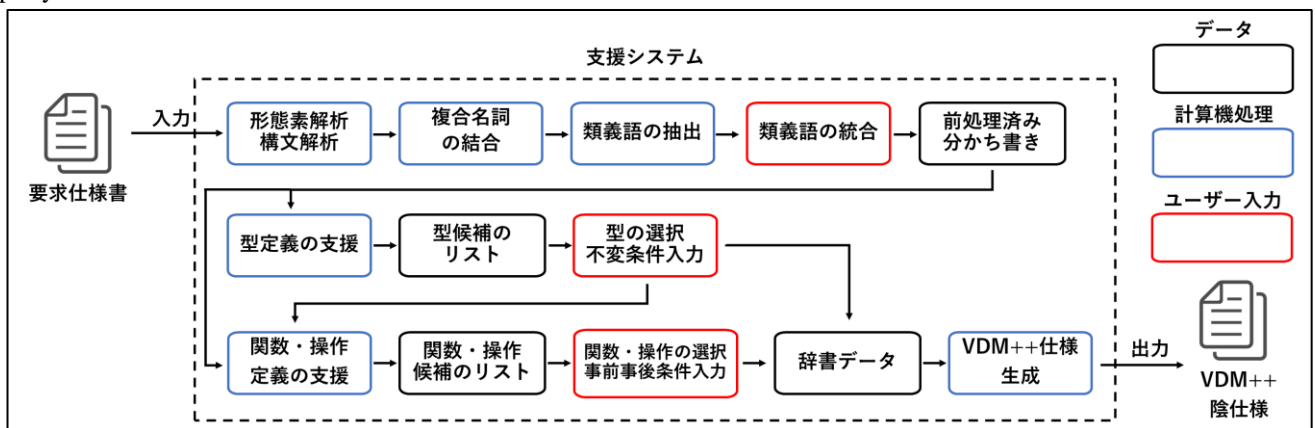


図2 変換支援手法全体の流れ

結合を行うことで、特定のドメイン知識によらない複合名詞抽出を行う。

3.5. 型定義の支援

本節で提案する手法では、名詞を列挙し、型になる可能性が高いと予想される名詞を優先的に表示することで、型定義を支援する。VDM++の一般的な作成手順で、「主に名詞から型を定義する」[7]とされているように、VDM++で型としてモデリングされるのは、要求仕様書中に含まれる名詞が主となる。そこで、本論文の変換支援手法では要求仕様書中に含まれる名詞および複合名詞を自動で抽出し、型の候補とする。また、型の候補は抽出された名詞および複合名詞全てとしている。全てを候補とすることで、作成者が確認すべき候補の数が多くなるというデメリットはあるが、本論文では、モデル化すべき単語の取りこぼしを減らすことを重視している。また、要求仕様書は専門性の高い文書であり、要求仕様書に含まれる用語の意味を正確に解釈することは困難である。そのため、最終的な判断は要求仕様作成者や対象ドメインの専門家のような人間によって行うほうが、正確な仕様記述が可能であると考えられる。しかし、型の候補を無作為に提示すると、作成者が必要な単語を探索する際の効率が悪いと、型になる可能性が高いと考えられる単語が上位になるように並べ替えることで、作成者は目的の単語を早期発見できるようになり、用語探索の負担が軽減される。

型になる可能性が高い単語について、専門用語のような文書の特徴が現れる単語は型になりやすく、多くの文書に含まれる一般的な名詞は型になりにくい。という仮説のもと、型候補の並べ替えに関して2つの実験を行う。

3.6. 関数・操作定義の支援

本節で提案する手法では、動詞を列挙し、関数・操作になる可能性が高いと予想される動詞を優先的に表示することで、関数・操作の定義を支援する。VDM++の一般的な作成手順で、「主に述語(動詞)から、関数・操作のインターフェース(シグネチャ)を記述する」[7]とされているように、VDM++で関数・操作としてモデリングされるのは、要求仕様書中に含まれる動詞が主となる。そこで、本論文の変換支援手法では要求仕様書中に含まれる動詞を自動で抽出し、関数・操作の候補とする。前節の型候補の並び替えと同様に優先的に提示する候補を持つ要素を考え、動詞が影響を及ぼす対象である目的語の関係を注目した。目的語は文章中の「何を、誰を」に当たる要

素である。これを VDM++に置き換えて考えると、操作はインスタンス変数を変える可能性がある操作であるため、影響を及ぼす対象が存在していると考えられる。また、検証用 VDM++仕様では、型名とインスタンス変数名が共通している定義が多かった。そこで、すでに定義されている型を目的語とする動詞を抽出することで、操作にあたる動詞を抽出できると仮定し、実験を行う。動詞の抽出に際し、抽出した動詞を形態素解析辞書の見出し語に変換する(レンマ化)。これにより、活用形を終止形に統一する。型を目的語とする動詞を抽出できたが、前述の仮定は操作の定義に対するものであり、インスタンス変数へアクセスしない関数の定義には対応できていないと考えられる。しかし、型候補の抽出と同様に、モデル化すべき単語の取りこぼしを減らすことを重視するため、型を目的語とするもの以外の動詞も関数の候補として提示する。

4. 変換支援手法の実現

4.1. 複合名詞の結合

複合名詞の抽出手法は複数あるが、本論文では2つの手法について実験を行った結果を示す。はじめに、仮説検定を用いたコロケーション抽出について実験を行った。コロケーションとは、物事について言及するため、慣用的に使われる2つ以上の単語からなる表現である。コロケーション抽出を行う統計的な手法として仮説検定を用いた方法が複数あり、その中で最も利用されているのが、尤度比検定に基づく手法である[14]。仮説検定は、データの統計量に基づいて、ある仮説の正否を検証する手法であり、尤度比検定は統計量として2つのモデルの尤度の比を用いている。算出した対数尤度比が高いほど、共起する可能性が高いペアだと言える[15]。対数尤度比が高い単語のペアを結合した複合名詞を表1に示す。

表1 対数尤度比が上位20個の複合名詞

	複合名詞	対数尤度比			
1	保温行為	171.616	11	水位センサ	91.857
2	温度制御	168.492	12	タイムアップ	86.041
3	蓋センサ	146.218	13	表示窓	83.608
4	操作量	124.574	14	min0	72.379
5	100msec	121.324	15	ロック解除	70.159
6	制御方式	115.356	16	0sec	68.151
7	要求仕様	111.954	17	水位メータ	63.178
8	沸騰行為	109.519	18	カルキ抜き	59.430
9	操作パネル	106.863	19	タイムボタン	58.043
10	残り時間	105.416	20	msec以上	56.267

次に、機械学習を用いた手法として、機械学習ベースの構文解析ツールである GiNZA を用いた実験を行う。3.3 節の通り、GiNZA は Transformers モデルを用いて日本語 UD の品詞ラベルと係り受けのラベルを付与する。GiNZA によって付与される、日本語 UD 係り受けタグの中に、複合名詞を表す”compound”関係があり、最小単位に分割された単語のなかで、複合名詞を構成しているものに付与される。そこで、”compound”関係を付与された単語同士を結合させることで、複合名詞とする。入力データの記述には、改行を句点相当の文の区切りとして扱う箇条書きのような記述が存在するが、GiNZA の出力で改行は空白として処理されており、文の区切りが不明瞭になっている場合があった。そのため、GiNZA に入力するデータは改行(¥n)を句点に置き換えている。また、”compound”と分析された単語の中で、名詞化の「こと」を含む組み合わせおよび、「()」の中にあることで、複合名詞とされているサ変可能な名詞を例外として結合しないこととした。 ”compound”関係の単語を結合した結果は表 2 のとおりである。

表 2 GiNZA により抽出された複合名詞(一部)

複合名詞					
1	ハードウェア構成	11	水位メータ	21	要求仕様
2	ハードウェア要求仕様	12	温度制御行為	22	操作パネル
3	内部構成	13	沸騰行為	23	給湯口
4	操作パネル部	14	保温行為	24	ポット上部
5	操作要求仕様	15	温度制御方式	25	正面側
6	沸騰ボタン	16	エラー検知	26	貯水部
7	保温設定ボタン	17	状態遷移	27	満水センサ
8	解除ボタン	18	話題沸騰ポット状態遷移図	28	買い物袋
9	給湯ボタン	19	話題沸騰ポット状態遷移表	29	未使用時
10	タイマボタン	20	話題沸騰ポット	30	本体接合部

4.2. 型定義の支援手法

・TF-IDF を用いた並べ替え

TF-IDF とは文章に含まれる単語の重要性を表す指標である。TF(Term Frequency)は単語の出現頻度、IDF(Inverse Document Frequency)は逆文書頻度を意味している。つまり、同じ文章内で多く出現し、他の文章で出現していない単語の重要度が高くなる。「話題沸騰ポット要求仕様書」から名詞を抽出し、GiNZA による複合名詞の結合を行った型候補のリストの各単語に対し、TF-IDF を算出し、TF-IDF の高い順に並べ替えた結果を表 3 に示す。

表 3 TF-IDF による並べ替え結果

	型候補					
1	水温	11	時	...	385	通常
2	on	12	off	...	386	際
3	保温行為	13	コンセント	...	387	上
4	°C	14	タイマ	...	388	他
5	ヒータ	15	こと	...	389	度
6	沸騰行為	16	セット	...	390	数
7	温度	17	ブザー	...	391	方
8	ポット	18	水量	...	392	状況
9	蓋	19	タイマボタン	...	393	設計
10	表示	20	操作量	...	394	もの

・FLR を用いた並べ替え

FLR とは、中川らが提案した専門用語抽出手法[13]で用いられた重要度を表す指標である。中川らは、ある単語や複合語が持つ分野固有の概念への関連性の強さである、ターム性と、ある単語や複合語がコーパス内で安定して使用されている度合いを表すユニット性に注目し、これらの性質反映した重要度の指標として、FLR を提案した。「話題沸騰ポット要求仕様書」から名詞を抽出し、GiNZA による複合名詞の結合を行った型候補のリストの各単語に対し、FLR を算出し、FLR の高い順に並べ替えた結果を表 4 に示す。

表 4 FLR による並べ替え結果

	型候補					
1	沸騰行為	11	水位	...	385	塩素
2	温度	12	モード	...	386	殺菌
3	温度制御行為	13	on	...	387	上
4	保温行為	14	時	...	388	タイミング
5	水位センサ	15	タイマボタン	...	389	無意識
6	温度制御方式	16	表示	...	390	指示
7	ポット	17	蓋	...	391	目次
8	蓋センサ	18	タイマ	...	392	状態 (アイドル)
9	沸騰ボタン	19	ヒータ	...	393	コストダウン
10	状態	20	時間	...	394	dE

4.3. 関数・操作定義の支援手法

検証用 VDM++仕様で定義された型に対して変換支援手法を適用し、型を目的語とする動詞を抽出した結果を表 5 に示す。

表 5 型を目的語とする動詞(一部)

	型	操作候補			
1	水位	検出	11	保温行為	中止
2	水温	検出	12	給湯ボタン	押す
3	コンセント	差し込む	13	給湯	停止
4	コンセント	抜く	14	タイマボタン	押す
5	蓋	閉じる	15	タイマ	起動
6	水位	確認	16	タイマ	操作
7	沸騰行為	する	17	タイマ	リセット
8	蓋	閉める	18	タイマ	停止
9	蓋	開ける	19	ブザー	鳴らす
10	沸騰ボタン	押す	20	操作パネル	見る

動詞の抽出に際して、関数・操作として定義されないと考えられる動詞の除外を行う。係り受け関係 fixed で助動詞に接続されている動詞は、複合機能表現と呼ばれる、単語の組み合わせによって慣用的に用いられる表現である。例としては、「～である」の「ある」や「～における」の「おく」などが該当する。これらの動詞は、組み合わせの一部であり、動詞の意味を直接表現するものではないため、除外する。また、サ変可能な名詞と「する」が接続され、サ行変格活用の動詞になっている場合、動作は語幹となる名詞に表れるため、名詞部分を抽出する。「する」が動詞の場合単体では動作の内容を特定できないため、候補から除外するが、助詞の「を」と接続し「～をする」の形(話題沸騰ポット中に含まれる例としては「沸騰行為をする」など)になっているとき、「を」の直前の語が目的語となり、「する」が動詞になる。このときの「する」は何かしらの動作を起こすという意味で、目的語で動作の内容を特定できると考えられるため、「～をする」は一つの動詞として扱う。以上により、型を目的語とする動詞のリストとそれ以外の動詞から一部を除外したリストが得られる。ここで、さらに効率的に単語を探索できるよう、得られたリストに対し個別に並べ替えを行う。並べ替えは、要求仕様書中に頻繁に表れる動詞ほど関数・操作となる可能性が高いという仮定のもと、出現頻度の高い順に並べ替える。並べ替えを行った結果を表 6 に示す。

表 6 並べ替えを行った関数・操作候補のリスト

	関数・操作候補				
1	押す	11	リセット	…	125 求
2	表示	12	保つ	…	126 取り除く
3	鳴らす	13	OFF	…	127 合う
4	検出	14	点灯	…	128 止まる
5	停止	15	起動	…	129 計算
6	消灯	16	知る	…	130 代用
7	閉じる	17	沸騰行為をする	…	131 対応
8	検知	18	確認	…	132 断線
9	制御	19	中止	…	133 カルキ抜きをする
10	止める	20	差し込む	…	134 吊り下げる

5. 変換支援手法の評価・考察

5.1. 検証用 VDM++仕様データの作成

支援手法の検討、検証を目的とし、「話題沸騰ポット要求仕様書」の VDM++仕様を人手により作成した。また、VDM++仕様として抽出される単語は一意に定まらないため、研究室所属の学生 3 名に「話題沸騰ポット要求仕様書」の VDM++仕様作成を依頼することで、複数パターンの VDM++仕様を作成した。なお、作成を依頼した学生は、プログラミングの基礎知識や、ソフトウェア工学の基礎知識を有するが、VDM++によるモデリングの経験はなかったため、VDM++モデリングについて 90 分程度の簡単な講習を行った後に、話題沸騰ポット要求仕様書についてモデル作成を行った。作成者のスキルとモデリングにかかった時間を表 7 に示す。

表 7 VDM++作成者のスキルと作成時間

作成者	学年	形式手法に関する経験	作成時間
著者	専攻科 2年	<ul style="list-style-type: none"> UMLによるモデリングの講義 90分×16週を履修済み Alloyを用いた形式手法の講義 90分×16週を履修済み VDM++の3ヶ月程度の学習およびモデリング 	4時間
学生A	専攻科 1年	<ul style="list-style-type: none"> UMLによるモデリングの講義 90分×16週を履修済み モデル検査手法 NuSMV および Spin の20時間程度の学習 VDM++に関する90分程度の講習 VDM++モデリングの20時間程度の学習 	6.2時間
学生B	本科 4年	<ul style="list-style-type: none"> UMLによるモデリングの講義を履修中 VDM++に関する90分程度の講習 	2時間
学生C	本科 4年	<ul style="list-style-type: none"> UMLによるモデリングの講義を履修中 VDM++に関する90分程度の講習 	2.7時間

なお、VDM++仕様はサブクラスへの分割を行わず、「話題沸騰ポット要求仕様書クラス」としてモデリングを行った。また、作成する仕様は陰仕様までとした。作成した VDM++仕様の一例を表 8、表 9 に示す。

表 8 著者作成 VDM++(一部)

```

types
public コンセント = <差し込まれている>|<抜けている>;
public 満水センサ = <on>|<off>;
public 第n水位センサ = <on>|<off>;
public 蓋センサ = <on>|<off>;
public ロック = <ロック>|<ロック解除>;
public サーミスタ = real
  inv temp == -10 <= temp and temp <= 150;
public ヒータ用電源 = <on>|<off>;
  ...
public 保温ランプを点灯 : () ==> ()
  保温ランプを点灯() == is not yet specified
post 保温ランプ = <点灯> and 沸騰ランプ = <消灯>;

```

表9 学生 A 作成 VDM++(一部)

```

types
public t稼働状態 = <on>|<off>;
public t温度 = real
  inv w温度 == -10.0 <= w温度 and w温度 <= 150.0;
public t水位 = <state0>|<state1>|<state2>|<stat...
public tモード = <高温>|<節約>|<ミルク>;
public t操作パネル ::
  沸騰ボタン : t稼働状態
  給湯ボタン : t稼働状態
  保温設定ボタン : t稼働状態
  タイマボタン : t稼働状態
  ...
public トリガ_保温設定ボタン : () ==> ()
  トリガ_保温設定ボタン() == is not yet specified
pre i操作パネル.モード = <高温> or i操作パネル.モ
ード = <節約> or i操作パネル.モード = <ミルク>
post i操作パネル.モード = <高温> or i操作パネル.モ
ード = <節約> or i操作パネル.モード = <ミルク>;

```

5.2. 評価指標

本章で使用する評価指標について述べる. 複合名詞の結合手法に関して, 再現率(Recall)を用いて評価を行う. 再現率は検索や推薦, 2 値分類の機械学習などのタスクを評価する際に用いられる指標である. 再現率の最小値は 0.0, 最大値は 1.0 となり, 値が大きいほど, 正例と予測すべきもののうち, 正例と予測できたものの割合が大きいことになる. また, 再現率は一般に, 正例と予測したもののうち, 実際に正例であったものの割合を表す, 適合率(precision)とともに用いられる.

型候補の並べ替えについて, PR 曲線(Precision-Recall Curve)と平均適合率(AP: Average Precision)を用いて評価を行う. PR 曲線は, 再現率を縦軸に, 適合率を横軸に取ったグラフであり, 予測データ数をしきい値によって変動させ, 各データ数における再現率と適合率をプロットする. 再現率と適合率はトレードオフの関係になっており, しきい値を下げ予測データ数を増やすほど, 再現率は上昇し, 適合率は低下する. 再現率が上昇しても適合率の低下が少ないものが理想的であるため, 予測の精度が良いほど右上に曲線が膨らむ. また, しきい値ごとの適合率を加重平均した値である AP や PR 曲線の下側の面積を積分計算した AUC によって PR 曲線は評価できる.

5.3. 複合名詞の結合手法についての評価・考察

複合名詞の結合手法について妥当性の評価を行う. 予測データは 4.1 節の実験により結合させた複合名詞と結合しなかった名詞から重複を取り除いた単語のリストで

ある. また, 複合名詞を結合しなかった場合の, 最小単位に分割された名詞のみのリストについても検証する. 正解データは, 検証用 VDM++仕様で, 型として定義された単語を合わせて重複を取り除いたものである(表 10).

表 10 型の正解データ(一部)

	型			
1	満水センサ	11	タイマ	...
2	第 n 水位センサ	12	タイマ状態	38
3	蓋センサ	13	水位メータ	39
4	ロック	14	保温モード	40
5	サーミスタ	15	保温温度	41
6	ヒータ用電源	16	水量	42
7	ランプ	17	動作状態	43
8	インジケータセル	18	沸騰行為	44
9	ボタン	19	保温行為	45
10	ボタンタイマ	20	稼働状態	46

3.1 節の通り, 本手法はモデル化すべき単語の選択は作成者の判断に依存している. そのため本手法では, モデル化すべき単語の取りこぼしを減らすことを重視しており, 候補の中から実際にモデル化されなかった単語の数は重視しないため, 再現率のみを用いて評価を行う. 各手法によって抽出された単語のリストの再現率は表 11 のとおりである.

表 11 各手法の再現率

複合名詞の結合手法	再現率	TP	TP+FN
複合名詞の結合をしない	0.43	20	46
GiNZA	0.76	35	46
LLR上位10ペアを結合	0.52	24	46
LLR上位20ペアを結合	0.57	26	46
LLR上位30ペアを結合	0.59	27	46
...
LLR上位80ペアを結合	0.65	30	46
LLR上位90ペアを結合	0.65	30	46
LLR上位100ペアを結合	0.67	31	46

表 11 より, 複合名詞の結合をしない場合に対して, GiNZA と LLR 双方ともに再現率が高くなった. この結果から, 人間が型として抽出する単語は複合名詞が多いと言える. また, GiNZA によって結合した結果の再現率が最も高く, 人間が抽出する複合名詞に最も近い手法であると言える.

以上の結果をもとに, 複合名詞の結合に関して考察を述べる. 抽出できなかった単語は, [ボタンタイマ, タイマ状態, 保温モード, 保温温度, 動作状態, 稼働状態, ロック解除中ランプ, 制御ランプ, 制御状態, インジケータランプ, タイマ時間]である. これらの単語のうち, 保温温度以外の単語は要求仕様書中に含まれていない単語で, VDM++仕様作成中に人間の知識によって個人の理解しやすい用語が作られたためである. 例えば, 「話題沸

「沸騰ポット要求仕様書」には、沸騰させた湯を保温する際の仕様として、高温モードは98℃、節約モードは90℃、ミルクモードは60℃で保温するという仕様がある。これらをモデリングする過程で、現在のモードや、各温度を一つの型で管理するほうが適切だと考え、「保温モード」や「保温温度」といった単語が定義されたと考えられる。以上のように、本論文で用意したVDM++仕様は、関数・操作のシグネチャまで記述しているため、仕様記述中に洗練が進んだのである。現行の技術では、要求仕様書中に含まれない語彙を抽出することは困難であるが、上位、下位概念の学習やなど組み合わせることで、予測できる可能性があると考えられる。

5.4. 型候補並べ替えについての評価・考察

型候補並べ替え手法について妥当性を評価する。3.4節の実験で並べ替えを行った型候補のデータに関して、PR曲線をプロットし、APを算出する。TF-IDFによる並べ替えを行ったデータのPR曲線を図3、FLRによる並べ替えを行ったデータのPR曲線を図4に示す。また、図3のAPは0.36、図4のAPは0.36である。

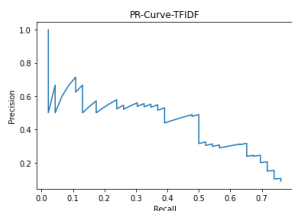


図3 TF-IDFのPR曲線

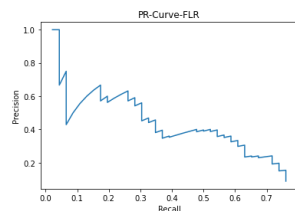


図4 FLRのPR曲線

以上の結果から、並べ替え手法の精度はあまり高くないと言える。並べ替えを行わない(要求仕様書中出现する順)場合のAPは0.21であるため、並べ替えを行わない場合と比較して精度は向上していると言える。しかし、図3、4より型として定義されない単語が上位に複数存在していることがわかる。

以上の結果をもとに、型候補並べ替えに関して考察を述べる。TF-IDFによる並べ替えでは、「ランプ」、「ボタン」、「センサ」といった単語が下位になった。これらの単語は「蓋センサ」や「沸騰ボタン」というように複合名詞となることで、頻度が下がったためだと考えられる。これらの具体的な要素を汎化したような単語の重要度を上げるには、WordNetのような単語の上位、下位概念を指標に取り入れる必要があると考える。また、上位になった定義されない単語の例として、「on」、「off」、「ポット」が挙げられる。「on」や「off」はセンサやボタンの状態を表す単語であり、VDM++では、型の値として定義される。「ポット」は「話題

沸騰ポット要求仕様書」のシステム全体を指す単語であり、クラスとして定義される。こうした、システムと等価なクラスである単語や状態を表す単語が高頻度で現れることが、TF-IDFの精度を下げる要因になったと考えられる。

FLRによる並べ替えでは、「ロック」や「サーミスタ」といった単語が下位になった。FLRは複合名詞の要素となっている単語の重要度が高くなるため、単独での出現が多い単語の順位が下がったと考えられる。また、TF-IDFと同様に、システムと等価なクラスである単語や、状態を表す単語が出現頻度によって重要度が高くなっていると言える。

次に、型になる可能性が高い名詞について考察を行う。専門用語のような文書の特徴が現れる単語は型になりやすいという予想は、抽出結果より概ね正しかったと考えられるが、仕様の中で最も特徴的な単語だと言える、モデル化する対象全体に相当する単語は型にならないという事実も見取れる。また、本論文で使用した指標は頻度を元に重要度を算出するものであり、仕様書中に一度だけ出てくる「第n水位センサ」のような単語の重要度が上がりづらいという問題もある。「温度」のような多くの文書で現れると考えられる単語が型として定義され、「ランプ」や「ボタン」のように具体的な要素を元に汎用的な単語が定義されたことから、一般的な名詞は型になりにくいという予想は、一概に正しいと言えないことが分かった。

本論文では、単語の専門性を型になる単語の特徴として扱ったが、実験により異なる特徴が見えてきたため、新たな指標を取り入れ、再定義する必要がある。解決案として、型候補の並べ替えに関する変換支援手法には、本論文の並び替えに用いた頻度に加え、上位、下位概念のような単語同士の関係性を表すオントロジーを指標に取り入れることが考えられる。単語同士の関係性を反映したデータとして、日本語WordNet[16]が公開されており、オントロジーを学習するための様々な研究が行われている[17]。単語同士の関係性を指標に取り入れることで、モデル化するシステムの構造が反映され、システム全体のクラスに相当する単語の提案や、具体的な要素をまとめ、汎化した単語を抽出することができれば、より高精度な型の抽出が可能になると考える。

また、「ポット」が抽出されたことから、抽出された名詞の中にクラスになる可能性がある単語が含まれていると言える。オブジェクト指向においてクラスと型の間に明確な違いはないため、システム全体を指す単語以外は、クラスと型どちらでも定義することが可能であり、基本的に作成者が判断する。VDM ToolsのVDM++開発手法ガイドラインでは、『用語集にある名詞は、そのモデルの目

的に関して“実質的”機能(読み／書きに加えるもの)をまったく持たないのならば、型としてモデル化されるべきである。』(九州大学 2016, p8)[18]とされている。この基準をもとにクラスと型を区別する手法を考察すると、関数・操作となる動詞の主語が機能を持っている名詞だと考えられる。そのため、関数定義の後ならばクラスとしての機能を持つ単語を区別する事ができ、将来的には洗練作業としてのクラス定義の支援が可能になると考える。

5.5. 関数・操作候補並べ替えについての評価・考察

関数・操作候補並べ替え手法について妥当性を評価する。予測データは 3.5 節の実験で目的語の関係をもつ動詞を抽出し、並べ替えを行った関数・操作候補のリストである。正解データは、検証用 VDM++仕様で、関数・操作として定義された単語を、4.3 節で行った動詞抽出のルールに合わせて整形を行い、重複を取り除いたものである。また、人手により作成したデータ中に「トリガ_保温設定ボタン」や「トリガ_タイマボタン」のように操作をモデリングした例がある。「ボタンが押された際に実行される操作」であると予測できるが、動詞を一意に定められないため、正解データから除外した。(表 12)。

表 12 関数・操作の正解データ(一部)

関数・操作				
1 差し込む	11	点灯
2 沸騰行為をする	12	点滅	30	解除
3 ロック解除	13	停止	31	検出
4 遷移	14	カルキ抜きをする	32	沸騰
5 切り替える	15	消灯	33	保温
6 ロック	16	中止	34	設定
7 給湯	17	検知	35	閉じる
8 止める	18	確認	36	加算
9 セット	19	判断	37	起動
10 リセット	20	鳴らす	38	終える

4.3 節で抽出され、並べ替えられたデータに関して、再現率は 0.82 となった。抽出できなかった単語は、[切り替える, 始動, タイムアップ, 接続, 切断, 要求, 到達]である。この内、「切り替える」と「切断」は「話題沸騰ポット要求仕様書」中に含まれない語彙であるため、抽出できなかった。しかし、「切断」はコンセントを切断する仕様だが、「抜く」といった候補は抽出できているため、定義することは可能である。「要求」や「到達」といったその他の単語は、「沸騰要求」や「沸点到達」といったイベントの名称となる複合名詞の一部として要求仕様書中に存在し、「要求する」や「到達する」という形で使われていなかったため、抽出できなかったと考えられる。これらの動詞は、形態素解析で最小単位に分割した際に、別個記録することで抽出

は可能であると考えられる。

関数・操作候補の並べ替え手法について妥当性を評価するため、並べ替えを行った関数・操作の候補のデータに関して、PR 曲線をプロットし、AP を算出する。関数・操作の候補の並べ替えを行ったデータの PR 曲線を図 5 に示す。

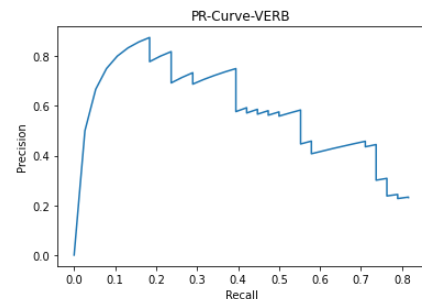


図 5 並べ替えデータの PR 曲線

図 5 の AP は 0.49 である。並べ替えを行わない場合(要求仕様書に出現する順)の AP が 0.25 であるため、並べ替えを行わない場合と比較して精度は向上したと言えるが、一位に不正解データが含まれており、改善の余地を残す結果となった。4.3 節で、すでに定義されている型を目的語とする動詞を抽出することで、操作にあたる動詞を抽出できると予想したが、型を目的語とする動詞の中で、実際に定義されていたのは 38 個中 23 個であったため、この予想は概ね正しかったと言える。

以上の結果をもとに、関数・操作候補並べ替えについて考察を行う。図 6 の通り、最も可能性が高いと判断された動詞の「押す」が正解データでは定義されなかった。これは、「話題沸騰ポット要求仕様書」内でボタンを押すという動作は頻出したが、ボタンを押すのは人間であり、ソフトウェアの動作ではないため、「押す」という語で定義されないと考えられる。そのため、主語が人間を表す単語(ユーザなど)となっている動詞の優先度を下げることで、改善が可能であると考えられる。しかし、日本語は主語や目的語が省略されるという特徴があり、実際には省略されている主語を予測するなどの課題があると考えられる。

6. まとめ

本論文では、形式手法導入コスト低減を目指すため、自然言語で記述された要求仕様書から VDM++仕様を作成する作業を、一部自動化により支援するアプローチの提案および、型定義、関数・操作定義に関する支援手法の実装を試みた。さらに、支援手法について評価・考察を行うことで知見の共有を行った。型定義の支援手法

に関して、人間の感覚に近い用語抽出を目指した、複合名詞結合手法の有効性を確認できた。一部抽出できなかった用語は洗練作業の中で作られたものだと考えられる。型として定義する用語を効率的に探索することを目的とした並べ替え手法は、型として定義される単語の特徴を重要度によって精度良く表すことはできず、改善の余地が残る結果となった。これらは、具体的な要素を汎化した単語やシステム全体を表す単語が実際には定義されないため精度が低下したと考えられる。関数・操作定義の支援手法に関しては、型を目的語とする動詞の優先度を上げることで実際に定義される単語を優先的に提示することができた。また、主語が人間である動詞は、実際に定義されづらいという知見を得ることができた。

参考文献

- [1] 独立行政法人 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, 「形式手法適用調査 調査報告書」, 2010-07-29.
- [2] 佐原 伸, 荒木 啓二郎「オブジェクト指向形式仕様記述言語 VDM++ 支援ツール VDMTools」, コンピュータソフトウェア, 24 巻 2 号, pp.14-20, 2007.
- [3] 「Overture Tool Formal Modelling in VDM」, <https://www.overturetool.org/>.
- [4] 千坂 優佑「VDM++仕様から C#コードを生成するツールの開発」, 情報処理学会全国大会講演論文集 78 巻 1 号, pp.363-364, 2016-03-10.
- [5] Yuma Yamano, Toshihiko Ando, Keishi Okamoto 「A Tool Generating a C# Code with Contracts of Code Contracts from a VDM++ Model with Conditions」, International Journal of Software Engineering (IJSE) Vol.8 No.3, pp.27-39, 2020.
- [6] 独立行政法人 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, 「厳密な仕様記述における形式手法成功事例」, 2013-01-25.
- [7] 独立行政法人 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, 「対称を如何にモデル化するか?」, 実務家のための形式手法シリーズ, 2013-03-28.
- [8] 大森洋一, 荒木啓二郎「自然言語による仕様記述の形式モデルへの変換を利用した品質向上に向けて」, 情報処理学会論文誌プログラミング (PRO) 3 巻 5 号, pp.18-28, 2010-12-10.
- [9] Yasuhiro Shigyo, Tetsuro Katayama 「Proposal of an Approach to Generate VDM++ Specifications from Natural Language Specification by Machine Learning」, 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), 2020.
- [10] 菅健将, 片山徹郎「VDM++仕様書におけるクラスとインスタンス変数定義の機械学習を用いた生成手法の提案」, ソフトウェアエンジニアリングシンポジウム 2022 論文集, pp.79-88, 2022-08-29.
- [11] SESSAME「話題沸騰ポット要求仕様書 (GOMA-1015 型) 第7版」, https://www.sesame.jp/workinggroup/WorkingGroup2/POT_Specification_v7.PDF, (最終アクセス 2022-01-11).
- [12] 松田 寛「GiNZA - Universal Dependencies による実用的日本語解析」, 自然言語処理 27 巻 3 号, pp.695-701, 2020-9-15.
- [13] 中川 裕志, 湯本 紘彰, 森 辰則「出現頻度と接続頻度に基づく専門用語抽出」, 自然言語処理 10 巻 1 号, pp.27-45, 2003-01-10.
- [14] Alice Zheng, Amanda Casari「機械学習のための特徴量エンジニアリング —その原理と Python による実践」, 株式会社ホクソエム訳, 2019-2-25.
- [15] Evert Stefan, 「The statistics of word cooccurrences : word pairs and collocations」OPUS - Online Publikationen der Universität Stuttgart, 2005.
- [16] 国立研究開発法人情報通信研究機構 (NICT) 「日本語 WordNet」, <https://bond-lab.github.io/wnja/>, (最終アクセス 2023-01-25).
- [17] 森田 武史, 山口 高平「オントロジー学習の現状と動向(<特集>オントロジーの進化と普及(前編))」, 人工知能学会誌 25 巻, pp. 354-365, 2010.
- [18] 九州大学「VDM++開発手法ガイドライン 1.0」, 2016, <http://fmvdm.org/>, (最終アクセス 2023-02-16)