

要求獲得におけるドメイン知識表現の役割

長田 晃[†] 小澤 大伍[†]
海谷 治彦[†] 海尻 賢二[†]

要求分析者にとって、ドメイン知識を得ることは要求獲得および仕様化に必須である。そのドメイン知識を安定的に得るためには何らかの形で明示化する必要があるが、明示化にはドメイン知識の表現形式が大きな役割を果たす。本論文では、要求分析の初心者がドメイン知識表現を参照しながら要求獲得を行う際の、ドメイン知識の表現形式によって受ける影響について実験を通して調査した結果を報告する。実験において、同じドメイン知識を、文章形式（用語辞書）と図形式（UML）の2つの表現形式で表し、要求獲得に使用した結果を比較した。結果として、用語辞書とUMLの違いに関して、ドメイン知識の表現が与える影響を確認した。

The Role of Domain Knowledge Representation in Requirements Elicitation

AKIRA OSADA,[†] DAIGO OZAWA,[†] HARUHIKO KAIYA[†]
and KENJI KAIJIRI[†]

It is important for a requirements analyst to acquire domain knowledge in order to elicit and specify requirements. To refer such domain knowledge stably, it should be represented in some kinds of form such as documents, diagrams and/or formulae. Thus, its representation gives influences on his/her requirements elicitation process and its outputs as well as its contents gives. In this paper, we explore effects by the representation itself through an experiment. In the experiment, we compared and analyzed outputs of requirements analysis each of which was developed with domain knowledge represented either in textual form (thesaurus) or in graphical form (UML). As a result, we confirmed the effects by the representations of domain knowledge, especially the differences between thesaurus and UML.

1. はじめに

顧客のニーズを把握するために、そのニーズに関連する業務やデータを知ることは要求獲得作業の重要な一部である。つまり要求獲得は、「顧客のビジネスに固有の知識」を必要とする。ここではそれらの知識をドメイン知識と呼び、顧客のビジネスを問題ドメインと呼ぶ。顧客との話し合いの中でドメイン知識を得ることはできるが、顧客は暗黙的な情報を明示しないことや、発生するかもしれない問題について明言しないことがある。さらには、そのようなドメイン知識の取得の仕方では、ドメイン知識はその場限りのものとなり、再利用ができず不安定なものとなる。そこで、ドメイン知識を明示化する技術が重要となる。

ドメイン知識を明示化する際には、内容だけでなく

表現形式についても十分に考慮しなければならない。なぜなら、異なる表現形式は、たとえ同じ内容を表現しているとしても、アナリストに異なる影響を与えるからである。たとえば、図書館の本の貸し出しの手順を、シナリオとして順序的に記述する場合と、アクティビティ図として記述する場合に、その内容はほとんど等しいと考えられる。しかし、アクティビティ図は各アクティビティを際立たせるとされるし、スームレーンが書かれていればアクタの認識についても影響を与えられる。しかしながら、設計などにおける表現形式についての研究はあるが⁽³⁾、要求獲得、特にドメイン知識の表現形式の適切さについての研究はあまりない。そして、設計ではより明示的な知識を扱うのに対して、要求獲得ではより暗黙的な知識を扱うことから、設計の手法をそのまま適用することは困難であると考えられる。

そこで我々は、ドメイン知識の表現形式の違いがドメイン知識の利用に与える影響について、実験を通し

[†] 信州大学大学院工学系研究科
Faculty of Textile Science and Technology, Shinshu University

て調査した結果を報告する。実験では、ドメイン知識の利用法の中でも、問題解決に必要な情報を抽出する部分に注目する。なぜならば、要求分析者はまず問題の把握を行わなければならない、問題を早く正確に識別する助けとして、ドメイン知識が利用されると思われるからである。よって、表現形式の適切さは、効果的な抽出ができたかどうかで判断する。この調査において明らかにしたい事柄は以下のことである。

- ドメイン知識の表現形式によって、抽出される知識には差が生じる。
- 工夫したドメイン知識表現は、より良い抽出が可能である。工夫とは、1方向からではなく、複数の側面から情報を表現することである。

2章では、関連研究について述べる。3章では、ドメイン知識と表現形式の関係を述べる。4章では、表現形式を測定する手法について考察する。5章では、表現形式による業務抽出の効果を測るためのメトリックスを定義する。6章では、表現形式の比較実験とその結果について考察する。さらに、実験についての妥当性についても考察する。最後に、現在の結果と今後の予定についてまとめる。

2. 関連研究

要求獲得においてドメイン知識は重要であるが、その表現方法についての研究は少ない。一方で、要求仕様書の表現についての研究は多い¹¹⁾。Katasonovらの研究⁸⁾において、仕様書の“視覚化”について考察しており、我々はこれを参考とした。

ここで典型的なドメイン知識表現のそれぞれについて概観する。UMLは、業務や組織に関する多くの種類の属性（構造や機能、振舞いの要素）を表現することができるため、代表的な表現形式の1つである^{5),12)}。よって、我々も実験においてUML記法を使用した。よく知られた表現形式の1つである“LEL”⁴⁾はクラス図に類似した表現である。“PAORE”⁹⁾と呼ばれる要求獲得技法は、機能の分解に主眼を置き、ドメイン知識を簡単な木構造で表現している。要求と現実世界との関係はドメイン知識を特徴づける重要な因子の1つであるが、“Problem Frames”⁶⁾の問題図は、ソフトウェア開発の問題となる構造を巧妙に表現している。しかしながら、その表現は問題分析者や要求分析者が彼らの問題を分析し、把握するためのものであって、ドメイン知識の利用者がドメインについてのより多くの情報を得るためには設計されてはいない。現在、多くの分野でオントロジが知識リソースとして使用されている。哲学的なものも含めオントロジには多くの定

義があるが、我々は単純に、オントロジとは“類語辞書 + 推論機構”⁷⁾であると考える。この定義を参考に、本論文では用語辞書を実験の中で用いている。

最後に、表現の比較のための研究について概観する。それを直接に述べた研究を見つけることはできなかったが、Katasonovらの研究⁸⁾の中では、「要求の表現形式を変更することにより、要求を多角的に分析できる」と述べられていた。また、ドメイン知識表現を使用した要求獲得は、学習活動の一種と見なすことができるため、学習結果を参照することができると思われる。

以前の研究において、機能の相関を行列により表現したドメイン知識を与え、そのドメイン知識の効果について確認する実験を行った。しかし被験者から、行列表現は必要な知識を見つけにくく、さらには思考を混乱させられた、との指摘を受けた。この指摘が、我々がドメイン知識の表現形式の役割を探索する理由の1つである。

3. ドメイン知識と表現形式の関係

本章では、ドメイン知識と表現形式について、図1を用いて説明する。我々は「問題ドメイン」の知識の中で、再利用可能な情報を「ドメイン知識」と見なす。すなわちドメイン知識は、明示的、暗黙的を問わず、問題ドメインについての情報で構成されている。明示的な情報は、たとえばマニュアルや仕様書などに含まれていると考えられる。しかし、暗黙的な情報はドメインエキスパートの頭の中にあるにすぎない。問題ドメインに詳しくない分析者は、それが具体的に表現されているときのみ、そのドメイン知識を参照することができる。よって、不慣れた分析者の分析が成功するために、表現形式は大変重要である。

要求分析のゴールの1つとして、顧客のかかえる問題やニーズに関連する業務の識別、つまり問題ドメイ

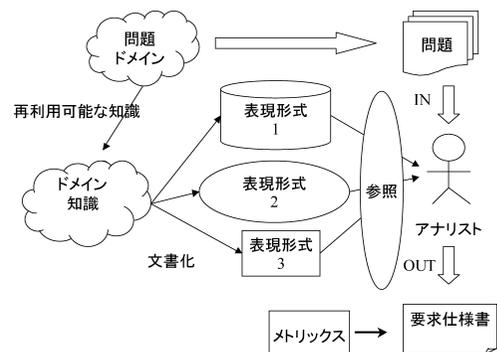


図1 ドメイン知識と表現形式の関係

Fig.1 Relationship between a problem domain and representations.

ンの一部を識別することがあげられる。要求仕様書はこのゴールを達成した結果に基づいて、つまり問題の業務を識別し仕様化を行った結果として、書かれるべきである。よって、表現されたドメイン知識の役割は、不慣れた分析者がこのゴールを達成するための役に立つことである。

図1に見られるように、ドメイン知識は複数の形式で表現することができる。しかし、どの表現形式がふさわしいかに関しては、問題ドメインの種類や顧客の意図、分析者の性質など、多くの因子が存在する。また、表現形式を直接かつ系統的に比較することは難しい。そこで、表現形式の違い以外は同じ条件になるよう調整して要求分析実験を行い、その結果を比較することとした。出力結果を利用して直接測定できないものを測り、メトリックスを用いて比較する手法は、合理的な手法であると考えられる。

4. 表現形式の測定

ドメイン知識の表現形式は、必要な情報の抽出に大きな影響を持つと考えられる。つまり、ふさわしい表現形式は情報抽出を正確かつ迅速に行う助けとなり、ふさわしくない表現形式は情報抽出において障害となるということである。よって、表現形式の効果について知っておく必要がある。しかし、ドメイン知識の表現形式の効果をそれ自身から直接測ることは難しい。

このような場合に、実際に使用すること、つまり実験を行いその結果を指針とすることは、単純で実際的な手法である。そこで、実験を通して、表現形式の効果を確認する手法について考察する。

表現形式に影響を与える代表的なパラメータとして、“問題ドメイン”、“分析者”、“問題自体”があげられる。このどれが違ってても、表現形式の効果が変化し、表現形式の適切さに大きく影響すると考えられる。問題ドメインについて考慮するときは、問題ドメインのカテゴリズのための研究¹⁰⁾を利用して、類似したドメインに対しては、同じ表現形式が有効であると仮定することができる。分析者について考慮するときは、複数の表現形式を提示し、分析者が好む表現形式を使用することができる。“問題”に関しては、問題自体の分類や副問題に分割して共通部分を再利用する提案⁶⁾を利用して、類似した問題には同じ表現形式を使用することができると考えられる。

よって我々は、複数の分析者が、あるドメインに関する複数種類の問題を解決する実験を行うことで、少なくともそのドメインに対しては、表現形式の効果を測ることができると考える。そしてその効果の評価は、

類似したドメインに対しても有効であると期待する。この実験に関して、表現形式以外にアウトプットに対する影響因子は、分析者の性質とインプットである。分析者の性質は、複数の分析者を用意することで、その影響を排除することができる。インプットに関しては、“Problem Frames”⁶⁾を参考に、“必要とされる振舞い”や“命令された振舞い”に関する情報を必要とする問題など、複数系統の問題で構成することで、表現形式の影響が特定の問題に依存しないように設定することができる。

5. メトリックス

本章では、ドメイン知識から部分集合を抽出する効果を測る実験のためのメトリックスを定義する。基本測定単位は単語とし、実験は入力された問題を解決する出力が得られるかどうかを測る。情報抽出は情報検索の一種であるので、情報検索のメトリックスである“再現率”と“精度”を参考に、良い抽出を以下のように定義する。

- 出力は正解の単語を含んでいる（再現率）。
- 出力は正解でない単語は含んでいない（精度）。

この“正解単語”はドメインエキスパートがドメイン知識と入力から選択したものとする。たとえば、「ドメイン知識を利用した要求獲得をしたい」という入力、および、ドメイン知識に「要求分析、分析者」などが含まれていれば、正解単語は「ドメイン、ドメイン知識、要求、要求獲得、分析、分析者」などである。もしこれらの単語が出力に含まれていないとしたら、それは悪い出力であると考えられる。

これらの抽出の出来を測定するために、4つのメトリックスを定義する。

- 網羅率

網羅率は出力結果が包含する正解単語の種類を数えるもので、再現率を参考にした指標である。網羅率が高いとき、出力結果は少なくとも正解単語について網羅していることを意味しており、分析者はドメイン知識から網羅的に知識を得ていると考えられる。たとえば、正解単語が100種類あり、出力結果に60種類が出現すれば、網羅率は60%である。

$$\text{網羅率} = \frac{|\text{出力の正解単語の種類}|}{|\text{正解単語の種類}|} * 100$$

- 要約率

出力に正解単語の数が多きことは重要なことであるが、出力全体も大きいことは、ふさわしくない。要約率は、出力全体の中の正解単語の割合を測る

もので、精度を参考にした指標である。要約率が高いとき、分析者はある単語が重要かどうかを識別して使用しており、出力結果には不必要な単語が少ないと考えられる。たとえば、出力が{論文, 投稿, 受理, 論文}であり正解単語が{論文}であるとき、要約率は $2/4 = 50\%$ である。

$$\text{要約率} = \frac{|\text{出力結果の正解単語の数}|}{|\text{出力結果の全単語数}|} * 100$$

● 重要度

基本的に重要単語は複数回出現すると考えられるため、たとえ出力結果に正解単語が網羅的に出現したとしても、1 回ずつだとしたら良い結果とはいえない。一方、2, 3 種類の正解単語が頻繁に出てきたとしても、良い結果とはいえない。ふさわしい出力結果とは、重要な単語が、網羅的かつ頻繁に出てくるものであると考えられる。重要度は網羅率と要約率を統合することで、全体的な品質を表現している。

$$\text{重要度} = \text{網羅率} * \text{要約率}$$

● 所要時間

要求獲得に限らないが、時間制限は最も大きな因子の 1 つである。そのため、実験における所要時間を計測する。分析者の性質の大きく依存するところではあるが、所要時間が短ければ、親和性が高いと思われる。

6. ケーススタディ

ケーススタディの目的は、知識抽出プロセスにおける表現形式の効果を検査することである。よって、典型的な 2 つの表現形式（文章形式と図形式）で表されたドメイン知識を用いた知識抽出実験により、効果を測る。図 2 は実験の概要である。実験の入力と出力を計測し、各表現形式を比較することで、効果を類推する。入力にはステークホルダからのニーズや問題をリスト化したものであり、“問題陳述書”と呼ぶ。出力はドメイン知識から抽出された業務の一部であり、“業

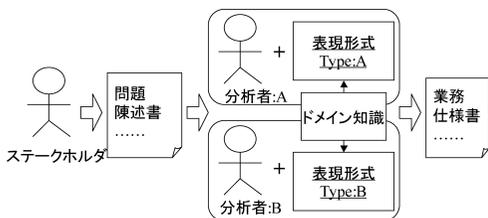


図 2 ドメイン知識表現の影響を測定するための実験
Fig. 2 Experiment to measure influence of domain knowledge representation.

務仕様書”と呼ぶ。2 つの表現形式を使用するために、被験者を 2 つのグループに分ける。

文章形式と図形式は 2 つの代表的な表現形式である。一般に、図形式は強調や補足が容易で、「1 つの図が 1,024 の単語に相当する」¹³⁾ などといわれている。図形式は情報を様々な側面から表現することができるため、我々は図形式を工夫した表現であると見なしている。しかしながら、図形式は文章形式より扱いが難しく、運用に時間がかかるといわれている。逆に、文章形式は親和性が高く、汎用的であり、運用が簡単だといわれている。もし親和性が高ければ短時間で理解することができると思われる。本実験で 2 種類の表現形式の効果を確認するために、1 章で述べた仮定を具体化し、以下の仮定をおく。

- (1) 文章形式と図形式の抽出結果には差が出る。
- (2) 図形式は文章形式より、より効果的に抽出が可能である。
- (3) 文章形式は図形式より親和性が高い。

6.1 ケーススタディのデザイン

ケーススタディのデザインについて説明する。

● ドメイン知識表現形式（入力形式）

文章形式として図 3 のような“用語辞書形式”を、図形式として図 4 のような“UML 図形式”を採用する。UML はモデリングの図形式としてはデファクトスタンダードであり、用語辞書は知識表現としてよく知られたものである。我々はこの 2 つの表現形式が、汎用的で使い勝手が良く、多くの専門知識を必要としない簡単なものであるため、ドメイン知識の表現として適切であると考えられる。

● 業務仕様書（出力形式）

業務仕様書は被験者が抽出の結果として書くものである。業務仕様書の例を図 5 に示す。問題ド

プログラム委員長:

プログラム委員会の長であり、会議のプログラムの決定に関して責任をもつ。主な仕事は、投稿された論文の管理とプログラム委員の選定と依頼、査読の依頼と査読結果の回収、プログラムの採択、不採択に関する最終的な決定である。

査読:

投稿された論文を査読者が評価し、プログラムに掲載するか否かの採否判定資料を作成する作業。

投稿された論文は、(ほとんどが複数の)査読者によって、採点が行われる。採点は、合格か不合格かの結果と、その理由も査読者によって付記される。

公正かつ適切な査読を行うため、当該論文の論じている分野に明るい研究者に査読者になってもらい、十分な時間を費やして査読することが望ましい。査読者は投稿後(すみや)に決定し、査読はプログラム委員会の開催日前には終了し、その結果をプログラム委員長に報告しなければならない。

図 3 用語辞書形式の一部

Fig. 3 A part of text dictionary representation.

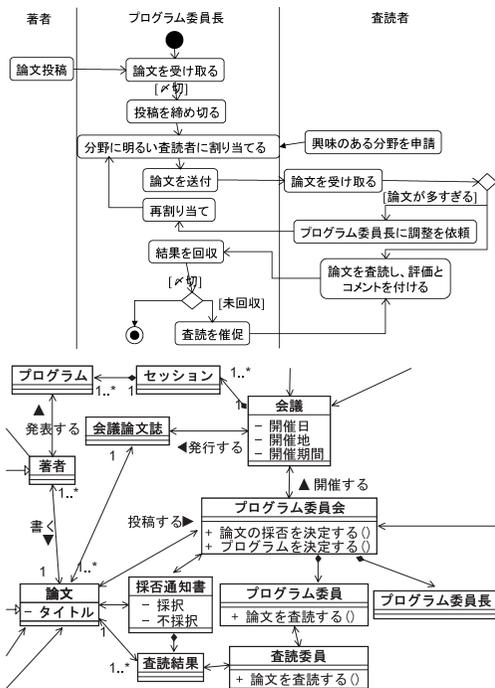


図 4 UML 図の一部

Fig. 4 A part of UML diagrams representation.

BSC:(論文受け取り)
 1.プログラム委員長は論文を受け取る.[4]
 ◇ 幾つかの形式で受け取り、保管する。
 ▶ 郵便で。
 ▶ PDFやMS-Wordなどの電子メールで。
 2.新規投稿が論文の更新かをチェックする.[3]
 ▶ 投稿者の申請。
 ▶ タイトルと著者から既投稿論文を照会する。
 3.受け取った論文を整理する。
 ▶ もし新規投稿であれば、保管する。
 ◇ タイトルと著者を記録。
 ◇ 連絡先を記録.[5]
 ◇ 投稿日時を記録。
 ▶ もし更新であれば、既投稿論文を更新する。
 ◇ 更新日時は必ず記録する。

図 5 業務仕様書の一部

Fig. 5 A part of business specification (output statement).

メインとしての業務の一部を抽出するという点から、シナリオ分析を参考に、実際の業務を階層的シナリオとして表現する。

● 問題陳述書

問題陳述書はステークホルダのかかえる問題を表現したものである。問題陳述書の例を図 6 に示す。これはステークホルダの問題を表現するという点から、一般的な自然言語で書くものとする。

| | |
|-----|------------------------------------|
| PS1 | 他の類似分野の会議との日程の衝突を避けるための情報が不足している。 |
| PS2 | プログラム委員に担当論文を割り振る作業が煩雑である。 |
| PS3 | 既に投稿している論文の更新作業に手間がかかる。 |
| PS4 | 投稿された論文を受け取り管理するのが大変。 |
| PS5 | 論文の投稿者に受け取った旨を連絡するのを忘れてしまうことがある。 |
| PS6 | 論文の切延長のプレッシャーが強くて困る。 |
| PS7 | 査読結果を返さない査読者がいるためプログラムの決定が遅延してしまう。 |
| PS8 | 論文の採否を決定する際、全部の論文を議論する時間的余裕がない。 |

図 6 問題陳述書

Fig. 6 Problem statements (input statement).

● ドメイン選択

ドメインの選択は重要な因子である。以前の研究において、ドメインがよく知られたものであったため、用意したドメイン知識表現は被験者にほとんど参照されなかった。被験者の知識の差によって影響を受けないようにするため、ドメインは被験者にとって不慣れなものであるべきである。

● 被験者の要件

被験者は UML 図（たとえばクラス図、ユースケース図やアクティビティ図）を知っているべきである。また、問題ドメインについて不慣れでなければならぬ。さらに、表現形式の効果が明瞭に反映されるために、独自の経験を持たない分析初心者が望ましい。

● 実験手順

- (1) ドメインを決定する。
- (2) ドメインエキスパートが“問題陳述書”を作成する。
- (3) ドメインエキスパートが、ドメイン知識を表現する用語辞書を作成する。
- (4) ドメインエキスパートが、UML 図を作成する。その際には内容の変化を避けるため用語辞書をベースとする。
- (5) 被験者を 2 つのグループ（用語辞書グループと UML グループ）に分け、要求獲得実験を行う。
- (6) ドメインエキスパートが正答を作成する。
- (7) ドメインエキスパートが、ドメイン知識表現と正答から重要かつ正解の単語を決定する。
- (8) 形態素解析を行い、出力結果を単語単位に分割する。その際には正解単語を形態素辞書に登録する。
- (9) 各メトリックスで測定する。

6.2 結 果

本学において、3回の講義を用いて実験を行った。講義参加者は単純に2つのグループ、文章形式の用語辞書を使用するグループと、図形式のUML図を使用するグループに分けた。そして参加者に対して、実験終了後にデータとしての使用許可を求めたところ、使用を許可した人数はUMLグループは14人、辞書グループは16人の計30人であった。

● 被験者

本学の情報工学科の3年生を被験者とした。彼らはUMLについてすでに学習済みであり、数人の学生は要求分析について興味を持っていた。

● ドメイン選択

“国際会議運営”業務をドメインとして選択した²⁾。これは多くの手順を持ち、被験者に不慣れなドメインであるため、適切であると判断した。そして、ドメインエキスパートは、プログラム委員を務めたことがある著者らの1人が務めた。

● ドメイン知識表現形式

－ Type A：用語辞書（文章形式）

これは用語の辞書であり、タイトルの語とその説明で構成されている。さらに説明文中に他のタイトル語へのリンクをつけた。実装にはWIKIを用いており、被験者はウェブブラウザを通して用語辞書を閲覧する。図3は用語辞書の一部である。これは“プログラム委員長”と“査読”の用語を説明しており、下線は他のタイトル語へのリンクである。

用語辞書形式の規模は以下に示す

タイトル語の数：32．1タイトルあたりの説明文の平均単語数：73．作成時間：7時間．

－ TypeB：UML図（図形式）

図4はUML図の一部である。この例のアクティビティ図は“査読”作業を表現している。ドメイン知識を表現するために使用したUML図はクラス図、アクティビティ図およびユースケース図である。クラス図は問題ドメインである国際会議運営についてのクラス、たとえば“論文、査読者、プログラム委員”などを表している。アクティビティ図は運営業務の各作業を表している。ユースケース図は会議運営とアクタの関係を表している。これらの図をJPG形式でウェブサーバに置き、被験者はウェブブラウザを通して閲覧する。

UML形式の規模は以下に示す。

クラス図：1，クラス：19．アクティビティ図：6，アクティビティ：70．ユースケース図：1，ユースケース：15．作成時間：16.5時間．

● 業務仕様書（出力）

業務仕様書は問題陳述書に書かれた顧客の問題に関連した実際の業務の一部を表すものである。当然ながらすべての業務が問題に関連しているわけではないため、関連した部分のみを表現する。図5は解答例として階層的シナリオによって書かれた業務仕様書の一部である。業務を表現する際にシナリオを用いることは、簡単で一般的な手法である。

● 手順

以下の手順で実験を行う。これは6.1節の実験手順の(5)、すなわち要求獲得実験についての詳細である。

(1) 事前学習と演習

被験者は要求獲得や分析に対して明確な指導を受けたことはなく、出力として使用する階層的シナリオの書き方について不慣れであると考えられた。よって、事前学習として、一般的な要求獲得や分析作業について解説を行い、簡単な演習により実験の流れと出力の書き方に慣れさせた。演習はドメイン資料を必要としないよう、被験者のよく知る図書館の業務を例題とした。全体として1.5時間をかけた。

(2) 本実験（業務分析）

“国際会議運営”を問題ドメインに設定し、2つのドメイン表現形式（UML図形式と用語辞書形式）の資料を用意した。それぞれを使う2つのグループに分け、業務分析を行った。問題陳述書は図6の8つの記述である。作業時間は2時間を基準とし、被験者の判断で提出を行い、提出時刻と開始時刻の差を各人の作業時間として測定した。

(3) 再試行

被験者に対して表現形式に関するアンケートをとるために、UMLグループと辞書グループの表現形式を入れ替えて、両方の表現形式を体験させた。そのうえで、各表現形式の利点と欠点、どちらをより好むかについてアンケートをとった。この試行に関しても業務分析を行っているが、学習効果が出ると思われるため、データとしては用

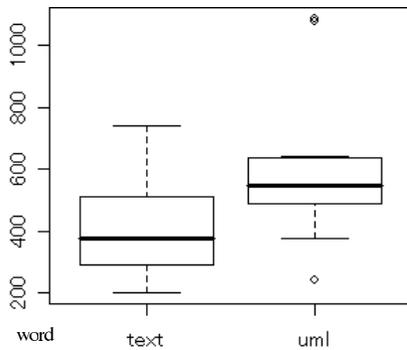


図 7 出力結果の全単語数

Fig. 7 Number of all words of output.

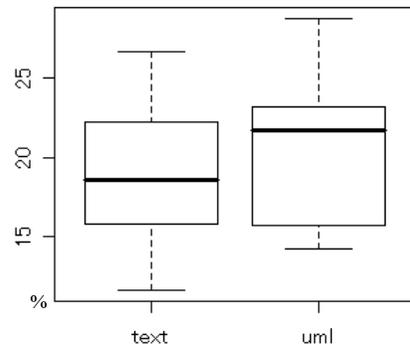


図 9 要約率

Fig. 9 Ratio of abridgement.

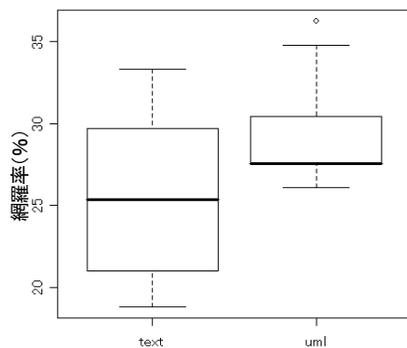


図 8 網羅率

Fig. 8 Ratio of comprehension.

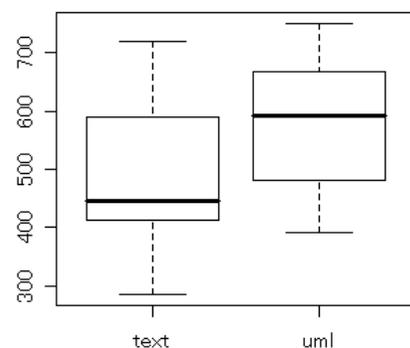


図 10 重要度

Fig. 10 Overall quality.

いない。

- 正解単語

69 個の正解単語を設定した。同義語はこの時点で統合し、正解単語内での同義語は存在しない。

- 正解単語のカウンターの仕方

正解単語は重複を許して、部分集合をカウントする。たとえば、単語“プログラム委員長”は、正解単語{プログラム, プログラム委員, プログラム委員長}をカウントする。このようなカウントの理由は、正解単語の複合語はその概念においても重複や含有があるという前提の下で、複数の知識が抽出されたと見なしているためである。

図 7, 図 8, 図 9, 図 10 から図 11 に、実験結果を提示する。全図を通して、横軸に text と uml をおき、測定結果をグループごとの対比として表した箱ひげ図である。縦軸は各図ごと左下に軸名を記載した。たとえば、図 7 は辞書グループと UML グループにおける、出力結果内の単語の数を箱ひげ図に表したものである。

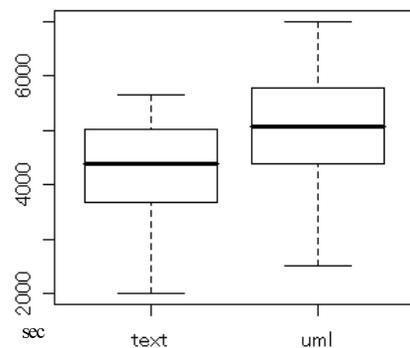


図 11 所要時間

Fig. 11 Working hours.

6.3 考 察

本章の冒頭で述べた 3 つの仮説について確認し、それぞれのメトリックスの結果に基づいて考察する。

(1) 文章形式と図形式において、実験結果には差が出る。

図 7 は出力の業務仕様書の規模である。同じシナリオ形式の出力結果を同じプロセスで処理し

単語を数えているため、この結果は両グループの規模の差としてふさわしいと考える。図 10 は抽出に関する出力結果の品質の差である。本実験は、表現形式以外はできる限り同様の設定とプロセスで行っているため、この差は両グループの差であると考えられる。図 11 は所要時間の差であり、辞書と UML に明確な違いがある。これらすべての結果にマン・ホイットニーの U 検定において有意水準 5% の片側検定で有意差があるため、この実験においては、文章形式と図形式に差が出たと考えられる。

- (2) 図形式は文章形式より、より効果的に抽出が可能である。

図 8 によると、UML グループは網羅率が高く、かつ、分布が狭いことが見てとれる。これは、UML 形式が網羅的な知識を補完し、内容の伝達に属人性が低いことを意味していると考えられる。また図 9 によると、UML グループの要約率が高いことが見てとれる。

この結果は、UML グループが効率的に抽出を行ったことを意味していると考えられる。UML 図、特にクラス図は、概念間のつながりを一目で把握できるような描写になっているため、網羅性が高くなったのではないかと考えられる。また、文章形式であれば補足説明として書くところを、記号形式で表すことができるため、より重要な概念のみが強調され、要約率が高くなったのだと思われる。さらにいえば、UML 図は概念が具体的で直感的に分かりやすく表現されているため、属人性の低い安定したドメイン知識を提供したと考えられる。なお、「テキストと UML のどちらが分かりやすかったか？」というアンケートに関して、73% が UML と答えている。我々はこれらの結果から、この実験においては図式形式がより効率的であったと判断してよいと考える。

- (3) 文章形式は図形式より親和性が高い。

図 11 は実験の所要時間の結果である。UML グループは辞書形式に比べ、より作業時間がかかっている。これは辞書形式が被験者にとってより簡単に親しみやすかったか、もしくは UML 形式が難しかったことを意味していると思われる。「資料として分かりやすかったか？」というアンケートに、良の評価を与えた被験者は、UML グループは 68% であったのに対し、用語辞書グループは 86% であった。さらに、「資料

として駄目だった」という評価をつけた人間は、用語辞書は 1 人もいなかったが、UML は 9% いた。このアンケート結果も、この実験において文章形式がより親しみやすかったことを裏付けていると思われる。

6.4 脅威と妥当性

ここで実験に関する脅威および妥当性について述べる。本実験は「ドメイン知識の表現形式の違いが、結果に影響を与える」という仮説を検証する実験である。この仮説に対して実験結果を示したが、その実験プロセス、実験内容および実験結果について、起こりうる脅威について考察することで、妥当性の確認とする。妥当性の分類は以下の 4 種の分類¹⁴⁾とする。

● 結論妥当性

処理と結果の間に正しい因果関係があり、妥当であるかどうか確認する。

－ 検定方法の選択

実験結果は、対応のない 2 グループであり、そのヒストグラムは正規分布していないものがある。よって、対応のないノンパラメトリック検定として、マン・ホイットニーの U 検定を用いた。なお、検定に用いたツールは“R”¹⁾である。

－ 被験者の無作為性

本実験は被験者が要求獲得初心者であるという前提をおいたうえでの実験と実験結果、および結論である。よって、被験者が要求獲得初心者の集合に含まれており、その中で無作為であれば問題はないと考える。

被験者は本学 3 年生のうち、ソフトウェア上・中流コースを選択した学生である。しかしながら、アンケートによると対象のドメインについて事前知っていた被験者は 1 人もおらず、また要求獲得作業も初めてであったことから、要求獲得初心者という前提において、被験者は無作為であると考えられる。

－ 分析方法の信頼性

データの分析方法は、文章を形態素解析にかけ、単語を抜き出して計測する手法である。これを 2 度行っても、結果が変わることはありえない。よって、分析の仕方に対する信頼性は高いといえる。

● 内容的妥当性

処理と結果の相関が意味のあるものであるとき、制御変数のみが結論に影響し、制御していない変数が影響を及ぼさないことを確認する。

－ 事前学習および学習効果の影響

事前学習の内容は、要求獲得やシナリオ分析、および図書館業務を例とした要求獲得作業である。ドメイン知識、およびドメイン知識表現については言及していない。また、事前に実験内容についての告知は行わず、資料は非公開であった。そしてデータを使用した実験は1回のみである。よって、被験者に学習効果やフィードバックによる影響はないと考える。

－ グループ分けの影響

30人の被験者のグループ分けにおいて、他講義の成績などはいっさい考慮していない。よって、一方のグループに質の高い被験者が集中した可能性は否定できない。しかしながら、2群に分ける際にその質が問題となるのは二極化している場合のみであり、一般的に被験者の能力はランダムに分散していると考えられる。そして、30個のランダムな変数を無規則で半分に分けた場合、その2グループに有意差が出る確率は5%ほどである¹⁾。よって、各被験者の能力はランダムに分散していると仮定し、グループ分けによる影響はないと考える。

－ 表記内容の同一性について

純粹に表現形式を比較するといった場合には、その内容を同じくすることが望ましい。しかしながら、文章をベースにUML図を作成する際に、文章のすべてをUML図に描くわけではなく、各々の表現形式の特徴を活かした形で書くこととなる。たとえば、図3において用語辞書に「査読」という項目があるが、その1行目は査読という作業の概要とそのゴールを明示している。一方、図4の「査読」のアクティビティ図では、全体のフローを把握しなければ査読の作業が理解できず、さらにはゴールが明示されていない。これは、フローを表現するというアクティビティ図の性質からくる制限である。しかしながら、アクティビティ図は「査読」における各アクタ、およびアクタ間のやりとりを明示しており、用語辞書では「プログラム委員長」「著者」「査読者」の3項目に記載されている査読に関する業務を集約し、1つの図に表現しているといえる。

このように、その表現の特性によって、1項

表1 表現形式の共通単語の割合

Table 1 Ratio of common words in both representations.

| | 辞書 | UML | 共通 |
|-----------|-----|-----|-----|
| 単語数 | 282 | 165 | 137 |
| 共通語の占める割合 | 49% | 83% | |

目や1つの図もしくは1画面内での内容には違いがある。本実験では、全体としてのコンテンツには違いがなくなるように努力したが、各部分での内容の同一性は保障できない。そのような各部分での内容の違いは、表現形式の特性の違いと考えられるため、別の影響因子ではなく、表現形式の違いとしての潜在因子の1つとしてみてよいと考える。

表1に表現形式の共通単語の割合を示し、表現形式全体における内容の同一性を確認する。この共通単語の導出方法を以下に示す。

- (1) 辞書とUMLからそれぞれに文章のみを抜き出す。
- (2) 形態素解析を行い、名詞と動詞のみを抜き出す。
- (3) 名詞のうち、数、および1文字で意味をなさない単語(たとえば“的”)を削除する。
- (4) 動詞は活用を考慮せず、すべて終止形へ変換する。
- (5) 同義語はより短い単語へ統一する(例:「しめきり、締切、締切り」は「締切」とする)。
- (6) 2群を比較し、共通単語を計測する。この結果から、本実験における表現形式には少なくとも半数の共通部分があるといえる。さらにUML図に存在する単語の83%は辞書に出ている。辞書における共通単語の含有率が低いが、これは辞書からUML図を書き起こす際に、図として記号化する部分があるため、文字列としては欠損していると考えられる。

我々は、UML図は辞書の部分集合となると考えていたが、本結果ではそうならなかった。表1によると、UML図の17%はUML図に固有の単語となっている。そこで、UML図の単語集合の中で、共通単語とならなかった単語を調べたところ、「~~切~~、タイトル」などが見られた。

「~~切~~」については、辞書では「締切」、UML図では「~~切~~」と表記していたためであった。

「タイトル」も同様で、辞書では「論文題目」と表記されていた。表記内容を同一にするよう努力したが、表記の揺れは制御できていなかったことを意味している。

しかしながら結局のところ、表記の揺らぎを考慮したとしても、8割が共通であるため、表記内容に明確な違いはないと考えられる。

－ 表現の提供法について

用語辞書とUML図はともに、個人のPCからウェブブラウザを通して参照しており、環境や作業としては同一である。また、用語辞書はWIKIを使用したHTMLファイルであり、UML図はJPGファイルであるため、どのようなウェブブラウザを使用したとしても、差異はないと考える。さらに、ウェブブラウザは一般的で難しい操作を必要としないため、ツールとしての影響はないと考える。ほかに、WIKIの検索機能が表現形式以外の因子としてあげられるが、今回使用した表現形式の「タイトルとタイトル間のリンクを持つ」という特徴は検索機能と類似した性質であるため、制御変数の範囲内であると考えられる。

ソーシャルな脅威として以下のものがある。

－ グループ交互の情報交換の影響

実験中における討論や相談は禁止していた。メールやWEBの使用についても禁止であると通告した。事前学習と本実験の間は1週間の空きがあるが、一般的に講義外での討論の可能性は低い。

－ 被験者のモチベーションの影響

複数の被験者を扱うことで、個々人のモチベーションについては吸収できると考える。

● 構造的妥当性

実験処理は原因を、実験結果は因果の結果を反映しているかどうかについて確認する。つまり、実験の設定や構造が、実世界の因果関係を適切に反映しているかどうかを確認する。

－ 実験の時間

実験の時間は8問で2時間である。これは著者ら自身が試行実験を行った際に、10問を回答するのにかった時間が50分であったため、被験者の能力およびドメイン表現資料を参照する時間などを加味し、当初16問を予定していたが半分の8問とした。よって、被験者らにとって十分な時間であったと思われる。また、結果的に2時間を超える被験者が

いなかったことも、実験時間が適切であったことを補強していると考えられる。

－ 資料の作成

用語辞書とUMLという2つのドメイン知識表現において、どちらか一方を重点的に作成してはいない。UMLの作成に時間がかかったのは、図を描くことに時間がかかったからである。しかしながら、用語辞書を先に作成し、UMLは2度目の作成であるため、作成者の慣れの影響を否定することはできない。

－ Mono-Operation Bias

本実験では、被験者、ドメイン知識表現形式、質問項目のそれぞれについて、複数の項目を用意している。よって、要素が単一である場合のバイアスはなく、現実世界の構造を反映していると考えられる。

－ Mono-Method Bias

本実験では、測定単位として単語のみを扱っている。そのため、単語に起因したバイアスが掛かっていることは否めない。

ソーシャルな脅威として以下のものがある。

－ 被験者が仮説を推理する。または評価を上げる気遣いをする。

仮説は「用語辞書とUML図において違いがある」ことであるが、被験者にはドメイン知識が2種類あることを明示していないため、他被験者と比較しなければ異なる表現形式があることを認識できない。よって、仮説を推理してバイアスをかけることは不可能であると考えられる。また、評価基準を公開していないため、評価を上げるために操作を加えることはないと考えられる。

● 外的妥当性

一般化に対する妥当性について確認する。

－ 実験の選択

被験者の選択として、今回は要求獲得初心者を対象とした。このため熟練アナリストなども含めた一般化はできない。

－ 実験の設定

業務の設定として、「国際会議の運営」のような、リソースを収集・管理する業務は一般によく見られる業務であり、現実的である。実験の構成として入出力を用いて計測対象を測る手法も一般的なものである。入力として、顧客がかなり具体的に問題点を列挙しているが、顧客が広範かつ具体的に問題点を列挙で

きるという設定は現実的ではないと思われる。しかし、顧客の問題点を抽出し分析する手法については本論文のスコープではなく、既存の研究を用いることで代替可能であると思われるため、一般化に影響はないと考える。

－ 実験の時期

本実験の前に、要求獲得に関する他の講義は存在していない。また、被験者が影響を受けようようなイベントも発生していない。よって、実験の時期は一般的であったと思われる。

7. おわりに

要求分析者にドメイン知識は必須であるが、効率的にドメイン知識を利用するためには、ドメイン知識を明示化する必要がある。その明示化の際には、ドメイン知識の内容だけでなく、その表現形式を十分考慮しなければならない。なぜならば、表現形式は要求獲得プロセスやその結果に大きな影響を与えるからである。よって本論文では、異なるドメイン表現形式の効果を実験を通して調査した。

実験においては、UML 図の 3 種類の図を用いて多方面からドメイン知識を表現した図表現と、用語辞書というシンプルな文章表現の 2 種類の表現形式を比較した。その知識抽出実験の結果は 2 つの表現形式に明確に違いが出るものであった。その内容は、図表現が効率的であったというものであり、我々は複数の側面から表現できるといった工夫した表現形式が、分析者にとってより良い抽出の助けとなると考える。これらの実験の結果から考えると、ドメイン知識表現は確かに、知識の抽出に影響を与えることが分かる。

4 章で述べたように、ドメイン知識の表現形式を測ることは、実験においても多くの因子が互いに影響しあうため非常に難しい。しかしながら、ステークホルダやドメイン知識作成者はドメインエキスパートであり、彼らは十分注意してドメイン知識や問題文を作成した。そのうえで結果に違いが生じ、さらにその結果は、表現形式が影響したと合理的に考えられる結果である。よって本実験において図表現が有効であったことは、因子の多様性により否定されるものではないと考える。そこで、同じく 4 章で述べたように、本結果は類似したドメインに対しても適用することが可能であると思われる。

我々はここで、2 つの表現形式で 1 つの問題ドメインを表現した実験を行った。このドメインについては UML が辞書よりふさわしいことを示唆していると思われるが、他の類似したドメインに対しても同様であ

るかについて確認するためには、さらに多くの実験が必要であると思われる。また効果の鋭敏性のため分析初心者を実験者として採用したため、分析初心者に限定したものとなっている。よって将来的にはパラメータを変えて実験を行う予定である。たとえば、1 つの表現形式で複数の問題ドメインを表現することや、被験者を 3 グループ（辞書、UML および両方など）に分けるなどである。ほかに、異なる表現形式、たとえば、表形式や形式的仕様記述などの式形式などが考えられる。また、単語での正当性については確認を行ったが、文章としての正当性の確認を行っていないため、それについての系統的な手法について研究する。さらに、図表現は作成に時間がかかるため、どのような図のどのような要素がより影響を与えるかについてを調査し、コストの削減をめざす。

本研究はドメイン知識を明示化するための第 1 歩として、ドメイン知識の表現形式を探求した。UML は運用に時間がかかるが、ドメイン知識を表現することにふさわしい表現形式であると考えられる。よって、UML を用いたドメイン知識を作成し、要求獲得におけるサポートに活用する予定である。ドメイン知識を得た分析者は、問題を正確に把握し、顧客への問合せを最小限に抑えられると期待する。

謝辞 本研究の一部は柏森情報科学振興財団 (K17 研 X 第 215 号)、および 2006 年度文部科学省科学研究費補助金 (基盤研究 (C) 課題番号 18500020) の支援を受けて遂行された。

参考文献

- 1) <http://www.r-project.org/>
- 2) <http://www.selab.is.ritsumei.ac.jp/~ohnishi/re/problem.html>
- 3) Arisholm, E., Briand, L.C., Hove, S.E. and Labiche, Y.: The impact of uml documentation on software maintenance: An experimental evaluation, *IEEE Trans. Softw. Eng.*, Vol.32, No.6, pp.365-381 (June 2006).
- 4) do Prado Leite, J.C.S. and Franco, A.P.M.: A Strategy for Conceptual Model Acquisition, *Proc. 1st IEEE International Symposium on Requirements Engineering*, pp.243-246 (1993).
- 5) Eriksson, H.-E. and Penker, M.: *Business Modeling with UML, Business Patterns at Work*, John Wiley & Sons (2000).
- 6) Jackson, M.: *Problem Frames, Analyzing and structuring software development problems*, Addison-Wesley (2000).
- 7) Kaiya, H. and Saeki, M.: Using Domain Ontology as Domain Knowledge for Requirements

- Elicitation, *Proc. RE'06*, (To appear), Minneapolis/St. Paul, Minnesota, USA (Sep. 2006).
- 8) Katasonov, A. and Sakkinen, M.: Requirements quality control: A unifying framework, *Requirements Engineering*, Vol.11, No.1, pp.42–57, Springer-Verlag London Ltd. (Mar. 2006).
 - 9) Kato, J., Saeki, M., Ohnishi, A., Nagata, M., Kaiya, H., Komiya, S., Yamamoto, S., Horai, H. and Watahiki, K.: PAORE: Package Oriented Requirements Elicitation, *Proc. 10th Asia-Pacific Software Engineering Conference (APSEC 2003)*, pp.17–26, Chiang Mai, Thailand, IEEE Computer Society Press (Dec. 2003).
 - 10) Maiden, N.A.M. and Hare, M.: Problem domain categories in requirements engineering, *Int. J. Human-Computer Studies*, 49(hc980206), pp.281–304 (1998).
 - 11) Solheim, H., Lillehagen, F., Petersen, S.A., Jorgensen, H. and Anastasiou, M.: Model-Driven Visual Requirements Engineering, *13th IEEE International Conference on Requirements Engineering (RE'05)*, pp.421–428 (2005).
 - 12) Tracz, W.: Domain analysis working group report: First international workshop on software reusability, *ACM SIGSOFT Software Engineering Notes*, Vol.17, No.3, pp.27–34 (July 1992).
 - 13) Wiegers, K.E.: *Software Requirements*, Microsoft Press (1999).
 - 14) Wohlin, C., Runeson, P., Host, M., Ohlsson, M.C., Regnell, B. and Wesslen, A.: *Experimentation in software engineering: An Introduction*, Kluwer Academic (2000).

(平成 18 年 11 月 28 日受付)

(平成 19 年 3 月 1 日採録)



長田 晃

2005 年信州大学大学院修士課程修了。同年より、信州大学大学院博士課程在学中。



小澤 大伍

2007 年信州大学大学院修士課程修了。



海谷 治彦 (正会員)

1994 年東京工業大学大学院博士課程修了, 博士 (工学)。1999 年より、信州大学工学部助教授。



海尻 賢二 (正会員)

1977 年大阪大学大学院博士課程修了, 博士 (工学)。同年信州大学工学部助手。1978 年信州大学工学部助教授。1995 年より、信州大学工学部教授。