

認知文法を用いた日本語から概念クラス図への変換法 Conversion Method from Japanese Sentence Specification to Conventional-Class-Diagram Using Cognitive Grammar

酒井孝真^{*1} 井田明男^{*2} 金田重郎^{*1*2}
Takamasa Sakai^{*1}, Akio Ida^{*2}, Shigeo Kaneda^{*1*2}

^{*1} 同志社大学大学院 理工学部

^{*2} 同志社大学

^{*1} Graduate School of Engineering, Doshisha University.

^{*2} Doshisha University.

要旨

要求分析では、仕様記述から概念クラス図を作成することで、分析対象を目に見える形に表記する。しかしながら、この過程がメソッド化されていないため、分析者に「迷い」の生じや、概念クラス図の「バラつき」といった問題が存在する。本稿では、上記の問題を解決するために、認知文法を用いた日本語仕様記述から概念クラス図への変換法を提案する。日本語の必須格である「を」「に」「で」を用いて仕様記述を単文化することで、日本語から概念クラス図へのリライトが可能となる。概念クラス図を一通り学習した学生を対象に実験を行い、手法適用前後で概念クラス図がどのように変化するか検証した。

1. はじめに

システム開発において、上流工程に位置する要求分析はシステム開発に大きく影響を与えている。現在では、数多くの分析手法が提案されており、代表的な分析手法ではOMG(Object Management Group)が管理するUML(Unified Modeling Language)の概念クラス図[1][2]や、Chenが提唱するER図(Entity-Relationship Model)[3]がある。これらの分析手法は図形を用いて分析対象を可視化する。分析者は可視化した図形を用いることで、利害関係者間の認識の差異を解消することが可能である。

UMLは世界的に用いられる分析手法であり、書籍も数多く存在する。日本でも多くの書籍が出版されている[1][2]。しかしながら、数多くのUMLに関する書籍において同一の傾向が伺えた。それは、概念クラス図の構成要素の説明や概念クラス図の使い方は説明されているが、仕様記述から概念クラス図への変換プロセスについては説明されていないことである。これは、あえて方法論化しないことで固定概念にとらわれず自由に概念クラス図を使う為かもしれない。しかしながら、分析者が初めてUMLを扱う初学者の場合、変換プロセスが不明確であるため概念クラス図を記述する際に「迷い」や「バラつき」が生じる問題がある。

本稿では、この問題を解決するために、仕様記述から概念クラス図を記述する変換法を提案する。具体的には、概念クラス図と日本語の構造を認知文法[4]の観点から分析し、概念クラス図と日本語の本質的な意味(コアイメージ)を模索する。コアイメージを用いることで、1対1で日本語から概念クラス図へリライトできる変換規則を提案する。

以下2章では本提案手法の概要、3章では本稿で用いる概念クラス図と日本語文法について認知文法の観点から分析した結果について述べる。4章では仕様記述から概念クラス図へリライトするための変換手順と規則について述べ、5章では実際に本手法を用いた適用実験と結果を示す。6章ではまとめについて述べる。

2. 提案手法の概要

著者らは、概念クラス図が英語の文型そのものであると主張している[5]。英語はコアイメージによって英単語の意味が1つ1つ明確に決まっている言語である。一方、日本語は、単語が多くの意味を持ち、雰囲気表現するあいまいな言語である。例として芭蕉の句「ふるいけや かわずとびこむ みずのおと」を挙げる。この句の英訳は数多く存在する[6]。なぜなら、日本語が雰囲気表現する言語であるため、カエルの数や響いた音の詳細が読み手によって解釈が異なるからである。同様の現象が要求分析を行う際に生じている。初学者が仕様記述から概念クラス図を作成する場合、仕様記述の本質を読み取れず様々な解釈を行ってしまい、結果として概念クラス図にバラつきが生じてしまう。

日本語と英語が根本的に異なるため日本語の仕様記述から英語の概念クラス図へ直接リライトすることは困難とされる。そこで認知文法を用いた間接的な変換法を提案する。認知文法を用いることで日本語が持つコアイメージの抜き出しが可能となる。抜き出したコアイメージを、英語の持つコアイメージと1対1で対応付けを行うことで機械的に仕様記述から概念クラス図へのリライトが可能となる。また、本手法を適用することで、仕様記述と概念クラス図の間にポイントが生成される。このポイントにより、概念クラス図を変更したとき仕様記述の変更箇所が容易に判断できる。

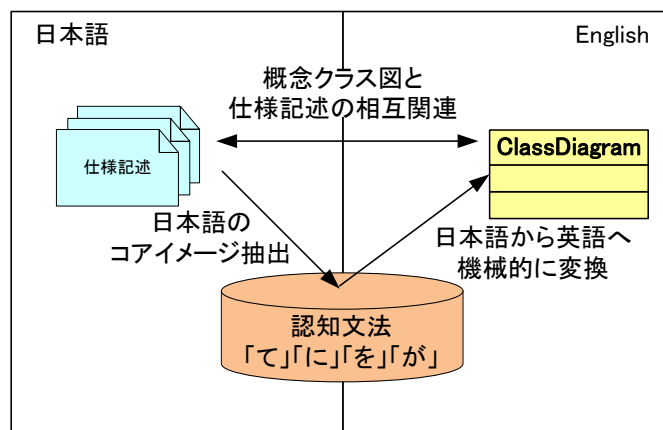


図 2.1 日本語から概念クラス図への変換法イメージ

3. 概念クラス図と日本語

3.1. 概念クラス図

「認知文法に基づくオブジェクト指向の理解」[5]に従えば、概念クラス図の構成要素が可算名詞・非可算名詞・動作動詞・状態動詞に対応しており(表 3.1)、概念クラス図が英語の5つの基本文型に一致する。本稿では5つの基本文型のうち、第三文型(S+V+O)と概念クラス図の対応に注目する。

概念クラス図のクラスと関連の関係は英語の第三文型(SVO)と一致するだけでなく、「S」と「O」に因果関係がある場合、つまり「V」が状態動詞をとる場合のみ第三文型と対応する。例えば「雇用主が社員を雇用する」の場合、雇用主がS、社員がO、雇用するがVとなる。ここで「S」と「O」の関係に着目する。社員から見た雇用主は操作主であり社員をコントロールすることが出来る。このように、支配関係ともいえるSがOを操作している状況を認知文法では因果関係[4]として表す。この因果関係をもたらす原因となるのがVの「雇用する」である。Vが「叱る」や「ほめる」といった動作動詞をとる場合、SとOの間に因果関係がなくなり、概念クラス図にリライトする事が不可能となる。実際にクラス図には「叱る」や「ほめる」といった動作動詞がクラス図の関連名として記述されている例は存在せず、状態動詞である「雇用する」などが関連名となる。動作動詞はメソッドとして記述される。

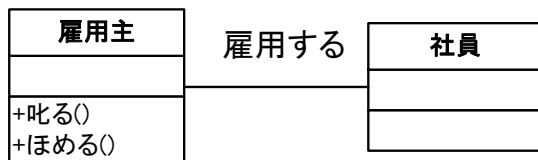


図 3.1 「雇用主が社員を雇用する」の概念クラス図

表 3.1 クラス図の構成要素[5]

認知文法における品詞	概念クラス図における要素
可算名詞	クラス名
非可算名詞	属性名
動作動詞	メソッド
状態動詞	関連名

動作動詞は一般的に関連名になることは無い。しかし、動作動詞が「S」と「O」の関連を表現する場合がある。それは、動作動詞によって「結果の残存」が生じる場合である。結果の残存とは動作動詞の影響により動作の前後で主語・目的語に状態変化が生じる場合である。「顧客が商品を購入する」を例に考える。購入するは動作動詞である。顧客が商品を持っていない状態から持っている状態に変化する。これを概念クラス図にあらわすと「もの—こと—もの」構造(図 3.2)[1]と同一の形になる。この種の動詞には、「購入する」「納入する」「発注する」などが存在する。

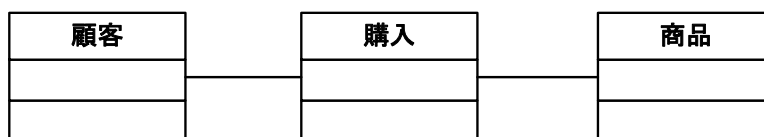


図 3.2 「結果の残存」による概念クラス図

3.2. 日本語

日本語のコアイメージを抽出するため認知文法の観点から日本語を分析する。日本語のコアイメージを抽出することで、日本語と概念クラス図が1対1で対応付けることが可能となる。森山新の「認知言語学から見た日本語格助詞の意味構造と習得」[7]に従えば、日本語の必須格と言える格は「ガ格」「ヲ格」「ニ格」「デ格」の4種である。格とは文節の区切りに用いる文字である。「私が学校に行く」の例では「私が」の「が」と「学校に」の「に」が格である。森山はさらに格を場合分けしており、「私が」の「私」は動作主、「学校に」の「学校」は移動先と定義している。森山が定義した必須格のコアイメージ(表 3.2)を参考に、著者らが定義した必須格のコアイメージを表 3.2 に示す。尚、今回は、日本語の文章ではなく日本語がSVOの型をとったとき「O」がどのように分けられるかを定義する。

表 3.2 著者らによる必須格のコアイメージ

ガ格	主体
ヲ格	対象
ニ格	対象, 場所, 時間, 目的
デ格	場所, 道具, 材料, 原因, 様態, 時間・金額

4. 日本語から概念クラス図への変換法則

日本語で記述された仕様書から概念クラス図へリライトする変換法則を提案する。具体的なステップを以下に示す。

STEP1 仕様記述の単文化

仕様記述を英語の第三文型(SVO)の形に単文化を行う。SVOの形とは「SがOをVする」、「SがOにVする」、「SがOでVする」の3種である。

STEP2 出現した名詞の抽象化

STEP1 で出現した可算名詞、非加算名詞の抽象化を行う。

STEP3 動作動詞、状態動詞の選択

STEP1 で記述した単文の動詞が動作動詞か状態動詞かを判断する。この際、目的語がとる必須格のイメージを記述するとSTEP4の作業がスムーズになる。

STEP4 概念クラス図へのリライト

単文を概念クラス図へリライトする。提案する変換規則(表4.1)にそってリライトすることで機械的な変換が可能である。単文化された文章が概念クラス図の構造と一致しているため一文一文が概念クラス図に対応する。

表 4.1 日本語と概念クラス図の変換規則

		状態動詞	動作動詞
ヲ格	対象	S-クラス, V-関連, O-クラス	S-クラス, O-S のメソッド
二格	対象	S-クラス, V-関連, O-クラス	S-クラス, V-S のメソッド
	場所	S-クラス, V-関連, O-クラス	S-クラス, V-S のメソッド
	時間	S-クラス, O-S の属性	S-クラス, V-S のメソッド, O-S の属性
	目的	状態動詞をとらない	S-クラス, O-S のメソッド
デ格	場所	S-クラス, V-関連, O-クラス	S-クラス, V-S のメソッド
	道具	S-クラス, O-S のメソッド	S-O のメソッド, O-クラス
	材料	S-クラス, O-S のメソッド	動作動詞をとらない
	原因	S-クラス, V-関連, O-クラス	動作動詞をとらない
	様態	状態動詞をとらない	S-クラス, V-S のメソッド, O-S の属性
	時間・金額	S-クラス, O-S の属性	動作動詞をとらない

以下に実際に変換法則を用いた例を示す。

【使用記述例】

XXさんは自動車を購入しました。三菱自動車の販売店で、パシエロの購入のため、販売員Aさんから総額300万円の見積りを得るが、金額が折り合わず。翌週、同じく、販売員Aさんが2回目の見積りで290万円を提案するが、まだ安くなると思ひ断る。3回目は販売員Bさんから250万円の見積りを得て、購入を決める。新車(川崎あ12-34)が6月末に納車され、代金250万円を支払った。

STEP1 仕様記述の単文化

1. XXさんが自動車を購入する
2. 販売員Aが見積もりを提案する
3. XXさんが見積もりで判断する
4. 販売員Aが値段を下げる
5. 販売員Bが見積もりを提案する
6. 販売員Bが値段を下げる
7. 自動車がナンバープレートを持つ

STEP2 出現した名詞の抽象化

「XX さん」は「顧客」である.

「販売員 A」, 「販売員 B」は「店員」である.

STEP3 動作動詞, 状態動詞の選択

- | | |
|--------------------|------------|
| 1. 顧客が自動車を購入する | ヲ格—対象—動作動詞 |
| 2. 店員が見積もりを提案する | ヲ格—対象—状態動詞 |
| 3. 顧客が見積もりで判断する | デ格—原因—状態動詞 |
| 4. 店員が値段を下げる | ヲ格—対象—動作動詞 |
| 5. 自動車がナンバープレートを持つ | ヲ格—対象—状態動詞 |

STEP4 概念クラス図へのリライト

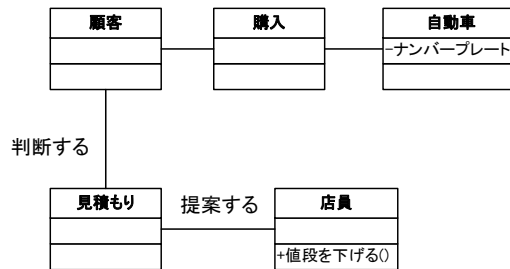


図 4.1 自動車購入の概念クラス図^{※2}

5. 適用実験

提案手法の有効性を検証するため, UML モデリングの初学者を対象に実験を行った. 被験者は大学講義で UML を学習した学生 5 人であり, 実際にシステム開発を行った経験は無い. 実験はテスト形式で概念クラス図を記述する問題を 2 問出題した. 被験者は, 初めに問題を自由に解いた後, 提案手法を用いて問題を解いた. 以下に問題文と実験結果の一部を示す(図 5.1, 5.2). 実験結果として, 手法適用前は分析者毎に様々な概念クラス図が記述された. 一方, 手法適用後は概念クラス図に大きな違い・バラつきは現れなかった.

【問題文】

A さん, B さん, C さん, D さん, らはあるオークションサイトの会員である
 A さんは, 今まで自分が使っていたデジカメとプリンタが不要になったので, それらを抱き合わせにして 10 月 7 日にオークションに出品することにした. 最低価格は 5000 円. それに対して B さんは, 10 月 8 日に 5000 円で入札, C さんは 10 月 9 日に 5100 円で入札, D さんは 10 月 10 日に 5200 円で入札. 結局出品期限である 10 月 10 日が過ぎても, D さんの入札価格を上回る入札がなかったため, D さんが落札した.

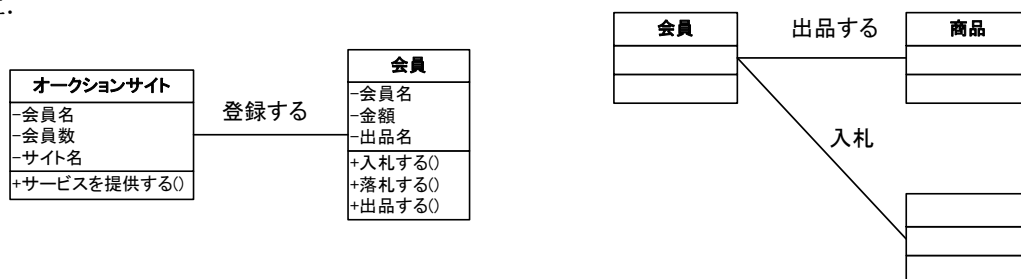


図 5.1 手法適用前の実験結果(一部)

^{※2} 図 4.1 を見ると「見積もり」と「自動車」に関連がひかれておらず関数従属を満たしていないことがわかる. これは, 基の様記述に「見積もり」と「自動車」の関係が記述されていないためである. 仕様記述から完璧な概念クラス図を作成することは困難であるため, 作成した概念クラス図をドメインと照らし合わせることで関数従属性の確認を行うことが重要である.

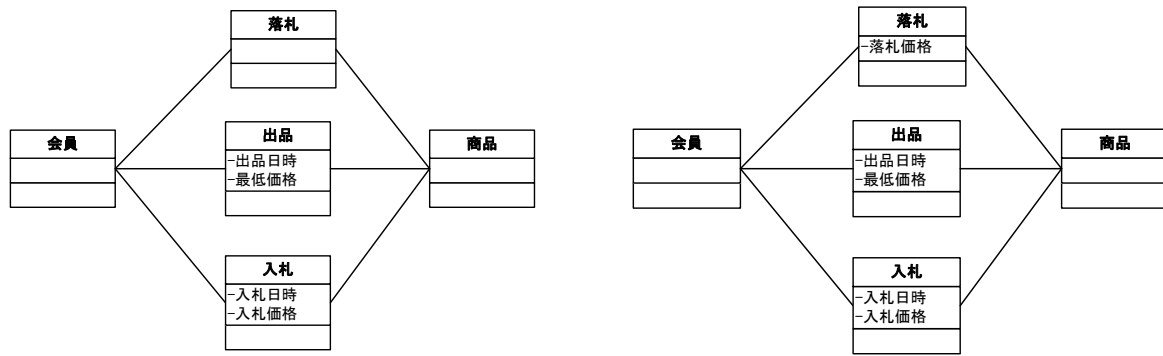


図 5.2 手法適用後の実験結果(一部)

6. まとめ

本稿では、認知文法を用いた日本語仕様記述から概念クラス図への変換法の提案を行った。適用実験を通じて「初学者に生じる分析時の迷いの原因」と「提案手法の有効性」について確認した。それぞれについて以下に述べる。

1 つ目は初学者に生じる分析時の迷いの原因について述べる。概念クラス図は英語の文型と一致しており、英語にはクラス間の関係を表す状態動詞も数多く存在する。一方で、日本語は文脈によって左右されるあいまいな言語であり、主語が省略されることが非常に多い。また、状態動詞も二種類しか存在しないため、日本語が英語圏で開発されたクラス図を用いるのに適した言語ではないといえる。そのため、日本人が概念クラス図を記述する際に日英翻訳と同等の作業を行う必要がある。実際に、クラス図を日々記述しているエンジニアは一度英語に翻訳したのちにクラス図を記述する人も少なくない。しかしながら、初学者は概念クラス図が英語という認識がないため、日本語の使用記述を概念クラス図に変換する際に何をクラスにしていかがわからない。おそらく「分析者の迷い」はこの部分にあるだろう。この迷いを解消するために、日本語を英語に訳すことが可能なフォーマットに変換することが必要である。この作業を本手法では「単文化」と呼ぶ。単文化を行うことで初学者でも迷いなく概念クラス図を記述することが可能となる。実験後、被験者にアンケートを行った結果、手法適用前は「何をクラスとして取り上げればいいのかわからない」「抽象化が難しい」といった意見が得られた。一方、手法適用後は「スムーズにクラス図を記述することが出来た」「迷いはなかった」といった意見が得られた。

2 つ目は単文化の有効性について述べる。システム開発を行う上で仕様書の変更やクラス図の訂正は数多く存在し、クラス図や仕様記述を変更した際にそれぞれに対応した箇所も変更する必要がある。概念クラス図を記述する際に仕様記述と対応付けを行っていないと、仕様記述とクラス図の対応関係がわからなくなってしまい、どこを訂正すればいいのかわからなくなる。しかし、本手法ではあらかじめ単文化を行っている。単文化を行うことで仕様記述と概念クラス図が一对一で対応付けられる。仕様記述と概念クラス図の対応を築くことで、仕様記述が変更された際にクラス図に影響する箇所がわかる。逆に、クラス図が変更された場合も仕様記述のどこを訂正すればよいかわかる。

最後に、今後の課題として、本提案手法では考慮していなかった「多重度」や「関数従属性」があげられる。

参考文献

- [1] 児玉公信, “UML モデリング入門 本質をとらえるシステム思考とモデリング心理学”, 日経 BP 社, 2008
- [2] 平澤章, “UML モデリングレッスン 21 の基礎パターンでわかる要求モデルのつくり方”, 日経 BP 社, 2008
- [3] P.P.Chen, “The Entity-Relationship- Toward Unified View of Data”, ACM Transactions on Database, Vol.1, No.1, pp.9-36, 1976
- [4] ジョン・R・テイラー, 瀬戸賢一, “認知文法のエッセンス”, 大修館書店, 2008
- [5] 金田重郎, 世古龍郎, “認知文法に基づくオブジェクト指向の理解”, 情報処理学会, 情報システムと社会環境研究会, SIG-IS, 2011
- [6] S.Hiroki, “One Hundred Frogs: From Renga to Haiku to English”, weatherhill, 1995
- [7] 森山新, “認知言語学から見た日本語格助詞の意味構造と習得 日本語教育を生かすために”, ひつじ書房, 2008
- [8] 岡智之, “日本語教育への認知言語学の応用～多義語, 特に格助詞を中心に～”, 東京学芸大学紀要, 総合教育科学系, 58, pp.467-481, 2007
- [9] 金田一晴彦, “日本語動詞のアスペクト”, むぎ書房, 1976