

概念体系を用いたオブジェクトの抽出手法

大山勝徳 山口猛 金子正人 武内惇 藤本洋

日本大学工学部情報工学科

オブジェクト指向分析において要求仕様を満たすオブジェクトモデルを作成するには分析者の経験知識が重要である。それぞれの人間が持つドメイン知識の構造が異なっている事から経験知識を系統的に獲得する事は困難である。本研究は、分析において経験に依存しないオブジェクトの抽出手法を確立する事を目的としている。オブジェクト抽出法を表すドメイン知識とドメイン知識を扱うための手続きをオントロジーによって明確化する。本稿ではオブジェクトの抽出に用いるドメイン知識の構築法とオブジェクトの抽出基準について論じる。

Object Extraction Procedure using Ontology

Katsunori Oyama Takeshi Yamaguchi Masato Kaneko Atsushi Takeuchi Hiroshi Fujimoto

Department of Computer Science College of Engineering Nihon University

Knowledge is important to make an appropriate model that satisfies required specification. It is difficult to show conceptual model of each one systematically. Thus the approach of acquiring knowledge is different. In our research it designates that establishing the object extraction Procedure does not need experience. For that, we assure that domain knowledge and procedures for object extraction is clear statement by ontology. We discuss construction of the domain model with a require specification description of controlled language, and clear statement conversion of the standard in object extracting.

1. はじめに

オブジェクト指向分析で良いオブジェクトモデルを作成するには、いろいろな視点から判断ができるようにするための経験知識が必要である。判断のポイントとなるのはオブジェクトを抽出する時に何に注目するかである。一般的にオブジェクトの抽出方法は一つに定まらない。これは人がそれぞれ持っているドメイン知識を含んだ概念モデル(図1)が異なるためである。

経験に依存したオブジェクトの抽出であると、作成されたモデルによるシステム化の意図が特に素人にとっては理解できないという問題がある。良いシステム開発を行うために、誰がオブジェクトの抽出をしても同じオブジェクトモデルに辿り着けるような方法が望ましい。そのような方法は開発者間だけでなく、ユーザとの対話においてシステム化の意図を相互に理解しやすくする。ユーザがシステム化の意図を早く理解できるようにすることにより、要件変更は早期化し、開発工程の後戻りの減少につながると思う。

例えば設計でデザインパターンを用いた時、動的モデル(さらにはコード)が記述しやすい(オブジェクトの分割性が高い)。

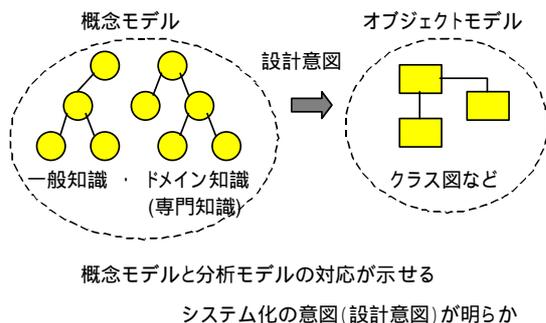


図 1 概念モデルとオブジェクトモデル

本研究では経験に依存しないオブジェクト抽出手法を確立する事を目的としている。そのためオブジェクトの抽出過程で、「どの知識」を使用し「どんな手続き」を実行したかを明らかにするメカニズムが必要である。オントロジーは知識マネジメントの立場から、知識の共有と再利用を実現するための技術である。そこで本稿ではオブジェクト抽出のための、オントロジーによるドメイン知識の構築法とオブジェクトの抽出基準の構成方法について論じる。

2. 研究の背景と位置付け

オブジェクトの抽出法に関連する研究では、自然言語解析結果から名詞に注目し、または動詞によって派生する関係性に注目してオブジェクトモデルのモデル要素を取り出す方法が示されている[1][2]。さらに、オブジェクトモデルを自動生成する研究[3]では要求文から生成された意味フレームと設定された分析ルールや調整ルールによりオブジェクトを生成している。しかし意味解析からは原文に忠実なモデルができるが、良いクラス図を描くためにはドメイン知識の活用が必要とされる。

オントロジーは対象概念を眺める視点を明示化する技術である。オントロジー技術はドメイン知識を体系化するので経験知識を上手く表現する有効な手段である。その応用例であるオントロジーを用いた通信サービス要求仕様記述の研究では、意味フレームに対して実行手順を表すタスクオントロジーを適用して、要求仕様書を検証する[4]。

オブジェクト抽出に関連するものとしてデザインパターン等、いろいろなパターン[5]があるが、これは経験から構築した事例を簡潔に表現したものといえる。パターンを使用する知識やその意図は人によって異なるため、使用理由が明確である事が重要である。また、パターンは断片的なドメイン知識であり、膨大であるため体系化が困難である。

本研究ではオブジェクト抽出へオントロジーを適用する事によりドメイン知識の体系化を図る。そのため、本研究は経験知識を明示化する所に特徴がある。経験知識の明示化により経験に依らないオブジェクト抽出を実現する。

3. オブジェクト抽出の課題

OMT[6]や Shlaer/Mellor 法[7]に見られるオブジェクトの抽出は主にオブジェクトのデータや状態に着目する。そのため一般的にオブジェクトは要求仕様やシナリオに現れる名詞からオブジェクトを判別している。また、RDD手法等ではシナリオから責務を抽出し、オブジェクトを判別する。本研究では Shlaer/Mellor 法を規範にする。この方法論においてオブジェクトの捉え方は現象(有形物・役割・出来事・相互作用・仕様など)ごとに大別されている。

オブジェクトの抽出に用いられる経験知識(概念モデル)は暗黙知であり人によって異なるため、開発者間や開発者・ユーザ間でオブジェクトを抽出した意図を相互に理解することが難しい。

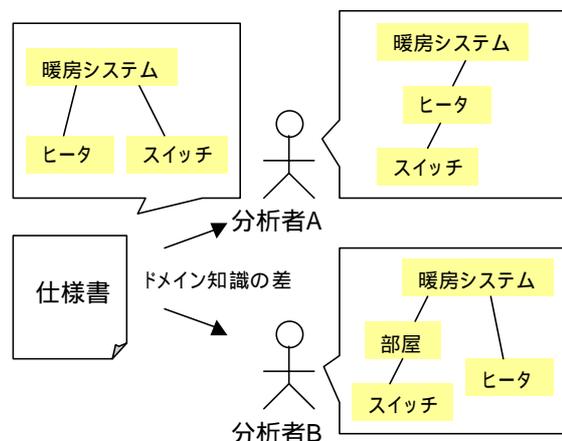


図 2 解釈の多様性(スイッチは部屋にあるのかそれともヒータにあるのか)

例として「暖房システムはスイッチで起動し、そのときヒータは部屋を暖める。」という文を考える。分析者 A がスイッチは直接ヒータにあるとし、分析者 B がスイッチは部屋の壁にあるとすると、暖房システムの規模や操作方法について認識の違いが生じている。また、仕様についての議論が不完全で元々の意味が部屋の中ではなく外に暖房システムの端末があり、そこにスイッチがあるとしたらユーザの要求を満たさず開発作業の後戻りが発生する（図 2）。

経験に依存しないオブジェクト抽出の方法を定式化するため、オントロジーの構成を図る（図 3）。すなわち、要求仕様に共通な概念の仕様を SSCJ（制限された要求仕様）オントロジー、オブジェクトに共通する概念（属性・関係や振舞い等）の仕様をオブジェクトオントロジーとし、両者のマッピング（抽出基準・手続きを表す）を明らかにする事により概念モデルとの対応を示せるようにするというものである。

4 . オントロジーの構成

4 . 1 定義

オントロジーは、工学の分野においては「暗黙的な知識を明示化したもの」^[8]、「概念化の

仕様」^[9]を意味する。すなわち、設計者がシステムを眺める視点についてモデル上で表現されているものをオントロジーと呼ぶ。従来、大規模なシステム開発ではシステムを構成する情報が膨大である。情報抽出を目的としたオントロジーは、情報抽出の対象の属性・構造などの意味情報と、さまざまな概念間の関係を公理的に定義することにより^[10]、膨大なシステム情報を扱いやすい形で記述する事が可能である。オントロジーを用いる事によって解決する主要な事項は合意を得る手段、再利用性・共有性、体系化・標準化である。

個々の人間は同一の対象物であっても視点が異なるためにそれぞれ異なった概念構造を持つ。しかしある部分では共通しているため、お互いにコミュニケーション（意図の伝達）をとる事ができる。つまり、それぞれの概念構造の共通部分を標準化したモデルとして明確にし、抽出することによりお互いの意図が解釈できるようになる。合意可能な標準化したモデルの例として UML がある。異なる方法論であっても UML を用いて表現する事により個々のモデルが理解可能となる事が挙げられる。

オントロジーには、問題解決そのものを表す「タスクオントロジー」と問題領域で表れる記述を定義する「ドメインオントロジー」の区別がある。

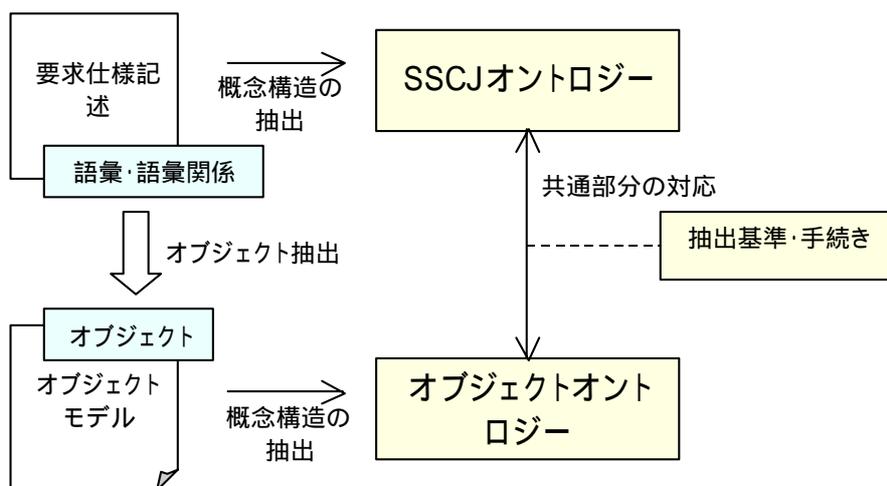


図 3 オントロジーを用いたオブジェクト抽出法の構成

タスクオントロジーは、問題の定義や問題解決能力、与えられる入力の規定や仮定、評価基準、問題解決のプロセス(方法・手順)などを表わす語彙を定義するものである。ドメインオントロジーは、問題解決の対象となる領域の記述に用いられる概念を定義するものである。

対象モデルに基づく問題解決システムの場合、ドメインオントロジーは中心的知識である対象モデルの規約を与える重要な役割を果たす[11]。以上の事から本稿で提案する SSCJ オントロジー・オブジェクトオントロジーはドメインオントロジーであると言える。

4.2 オントロジーの構成

このアプローチにおいては以下の事柄を検討する必要がある。

- 要求仕様書の概念構造を表現するための SSCJ オントロジー
- Shlaer/Mellor 法で捉えたオブジェクトの概念構造を記述するためのオブジェクトオントロジー
- 両者のドメイン知識のマッピングを表す抽出基準と手続き

4.2.1 SSCJ オントロジー

SSCJ オントロジーは、対象システムで実現しようとする機能を概念化する。以下の概念構造からなる。

- 目的：記述対象システムの機能の目的や意志
- 機能：発生するイベントに伴うデータの出入力と状態遷移
- 構造：上位・下位関係(親子関係)
- 制限：アルゴリズムやデータ構造・イベント発生に関する特性記述

このオントロジーが表すモデル(図4)ではシステムが果たす機能を表す「機能体」が最上位である。

4.2.2 オブジェクトオントロジー

オブジェクトオントロジーは、オブジェクトの属性・状態・オブジェクト間の関係など、モデル化に必要な構成要素を体系化することを目的とする。モデル化した時、オブジェクトは「オブジェクト名」「関係」「属性」「状態遷移」

で構成される(図5)。オブジェクトオントロジーはオブジェクトの静的モデルであるクラス図に対応している。抽出されるオブジェクトはオブジェクトの現象(有形物・役割・出来事・相互作用・仕様)によりモデルに表れる特徴が異なる。そのため、オブジェクトオントロジーは識別した現象ごとに特徴を表現する必要がある。

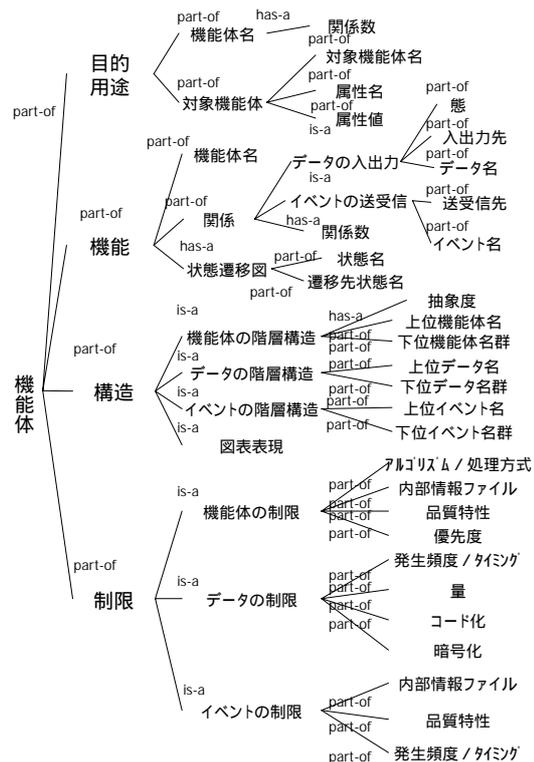


図4 SSCJ (制限された要求仕様) オントロジー

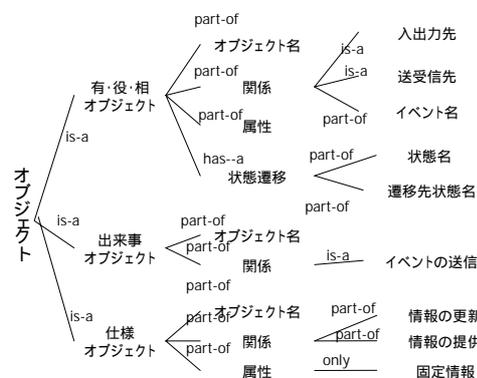


図5 オブジェクト (Shlaer/Mellor 法) オントロジー

4.2.3 オブジェクト抽出基準

SSCJ オントロジーとオブジェクトオントロジーの共通部分を明確にするため、オブジェクト抽出基準の定式化作業を行った。分析対象システムには「テニスボール封入システム」を使用した。その際に要求仕様書は自然言語で記述されるため語彙の不統一、係り受け等を解析するのに負担が強いられる。ここでは問題を簡単にするために要求仕様書には制限言語^[12]を用いる。以下に、オブジェクト抽出基準の定式化作業手順を示す。(図6参照)

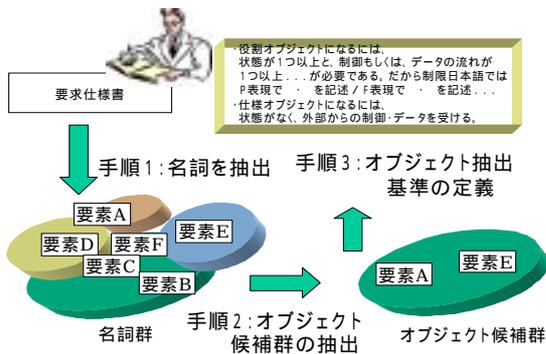


図6 オブジェクト抽出基準の定式化手順

手順1 名詞抽出

要求仕様書から名詞を取り出す。その結果、名詞の種類は約3000語の要求仕様記述に対して自然言語(係り受けが不定)では約160、制限言語では約120表れる。

手順2: オブジェクト候補群の抽出

テニスボール封入システムを手作業によって分析した結果を対象に、オブジェクトとして抽出された要素の名前と、手順1によって要求仕様書から取り出された名詞を比較する。手作業の分析結果と同一のものであればオブジェクトとなり得ると判断でき、その名詞をオブジェクト候補とする。

手順3: オブジェクト抽出判断基準の定義

オブジェクト候補群に分類した要素を取り出し、Shlaer/Mellor法で定義されている

オブジェクトの各現象(有形物・役割・出来事・相互作用・仕様)と比較する。このとき、SSCJ オントロジーの構成要素(「機能体」、「状態」、「属性」、「データ・制御の流れ」など)の組み合わせ条件をオブジェクトオントロジーと対応させて得る。この組み合わせ条件を「オブジェクト抽出基準」とする。

以上により作成されたオブジェクト候補リストを表1、オブジェクト抽出基準表を表2に示す。

表1 オブジェクト候補リストの一部

No.	抽出名	属性	label	出現数	判定	判断内容
1	商品	属性・制限	p02	5		
2	テニスボール	機能体	p03	10	O	機能体 有形物
3	1つ	属性・制限	p05	2		
4	缶	機能体	p06	20	O	機能体 有形物
5	充填	サ変名詞	p07	2		
6	テニスボール封入システム	システム構成要素	c01	3		
7	作業範囲	属性名	c02	1		
8	1つ	条件・状態	c03	2		
9	工場内	属性・制限	c04	1		

表2 オブジェクト抽出判断基準表の例

種類	有形物	
オブジェクト抽出判断基準		
< 制限言語記述 目的表現において >		
・目的表現に該当する機能体(動作主)がある。 (記述対象要素に動作目的がある)		
< 制限言語記述 機能表現において >		
・状態・環境が1つ以上記述されている ・他の機能体に対しイベント・データを送受信している		
< 条件式 >		
$f((P(f) \ F(f)) \quad d \cdot F(d) \quad s \cdot F(s) \quad 0(f))$		
P:目的表現	F:機能表現	0:オブジェクト
f:機能体	d:データ送受信	s:状態・環境

オブジェクト抽出基準表により「機能体」や「イベント・データ送受信」とのマッチングにより条件解析が図7のようになされ、「ボール供給機」が得られる。しかし複数のオブジェクト抽出基準表を用いるには、次節のオブジェクト抽出手順の定式化が必要となる。

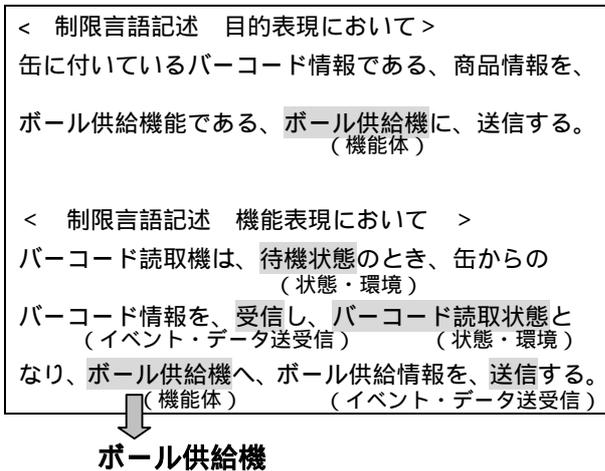


図7 オブジェクト抽出基準表によるオブジェクトの抽出例(ボール供給機)

4.2.4 オブジェクト抽出手順

オブジェクトの抽出を行うためには、オブジェクト抽出基準表を用いてオブジェクト抽出処理を行う。その時、オブジェクト抽出手順の定式化が必要となる。本稿で用いたオブジェクト抽出手順は以下の通り(図8参照)。

手順1 準備

要求仕様書を記述した後、名詞抽出をし、オブジェクト候補リストを作成する。

手順2 オブジェクト候補の分析

オブジェクト候補リストから、すべての機能体を分析するまで以下の処理を繰り返す。

仕様のオブジェクト抽出基準表で条件比較し、オブジェクトオントロジーの構成を満足すれば、オブジェクトと判断する。

仕様オブジェクトではない場合、出来事オブジェクト 相互作用オブジェクト 役割オブジェクト 有形物オブジェクトの順番にそれぞれオブジェクトの条件を

満足しているかを判断し、満足すれば、オブジェクトと判断する。

手順3 結果

全てのオブジェクト候補について手順2が完了しても条件を満足しなかった場合は、その機能体はオブジェクトではないと判断する。

また、全てのオブジェクト候補について手順2が完了済みのとき、オブジェクト抽出が終了する。

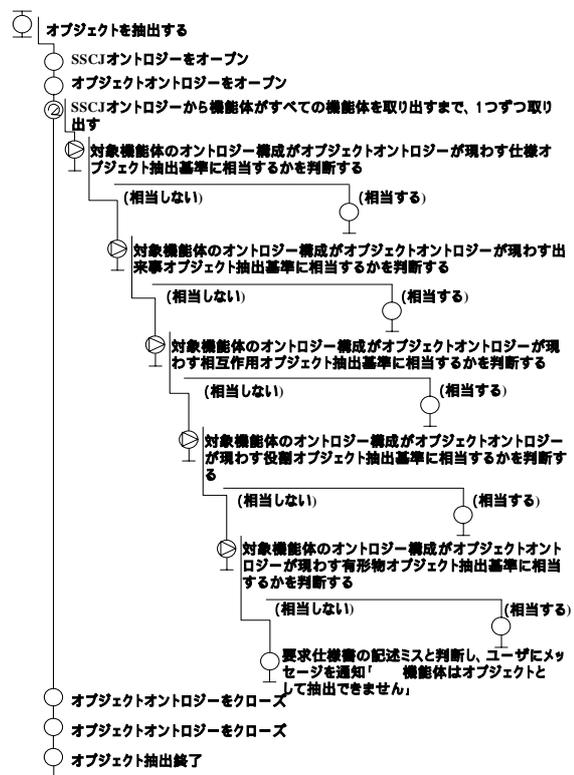


図8 オブジェクト抽出手順 HCP チャート

5. オブジェクト抽出実験

抽出基準を用い、誰でも同じオブジェクト抽出ができることを確認するため「テニスボール封入システム」を例に検証実験を行った。

オブジェクト抽出判断基準、オブジェクト抽出手順の作成者を含む4名で以下の同じルールを用いた場合のオブジェクト抽出結果の合致率を求め、また経験者の手作業による抽出結果との比較をした。

5.1 オブジェクト抽出結果の比較

表3に示すように、4人の検証者がそれぞれ同一のオブジェクト抽出基準表を用いた場合においては同一のオブジェクト抽出結果となった。今回の分析対象システムではオブジェクトの現象のうち、役割は確認されなかった。

表3 オブジェクト抽出結果(検証者4人)

	出現頻度	検証者A	検証者B	検証者C	検証者D
ボール貯蔵庫	8	有形物	有形物	有形物	有形物
缶搬送機	28	有形物	有形物	有形物	有形物
搬送回数カウンター	3	有形物	有形物	有形物	有形物
缶残量センサー	2	有形物	有形物	有形物	有形物
バーコード読取機	3	有形物	有形物	有形物	有形物
ボールセンサー	7	有形物	有形物	有形物	有形物
ふたセンサー	2	有形物	有形物	有形物	有形物
缶密閉機	16	有形物	有形物	有形物	有形物
制御パネル	29	相互作用	相互作用	相互作用	相互作用
ボール供給機	30	相互作用	相互作用	相互作用	相互作用
受信	21	出来事	出来事	出来事	出来事
出力	7	出来事	出来事	出来事	出来事
送信	27	出来事	出来事	出来事	出来事
缶	20	仕様	仕様	仕様	仕様
テニスボール	10	仕様	仕様	仕様	仕様

5.2 熟練者の抽出結果との比較

表4のオブジェクトリストはそれぞれの場合で生じた差異を表している。抽出1は前述の抽出基準を用いた結果を示し、抽出2は実際に熟練者によって抽出されたオブジェクトの例である。

主な差異は以下の通り。

- 実装レベルの語彙
- 汎化・特化
- 関係から得られるオブジェクト

抽出1の場合においては「受信」など実装レベルの語彙が多い。それに対し、抽出2は実装に近い「センサ」を除いている。他に、抽出1では汎化・特化が施されていないが、当然その

過程を設けなかったためだと言える。抽出2では「缶」に対し仕様を設け、特化をしている。そして、抽出2で見られる「ボール補充要求」は文章では直接表れない関係から判断して得られるオブジェクトの例である。

表4 オブジェクトリスト(太字は差異を表す)

抽出1(表1の抽出基準表)	抽出2(熟練者)
ボール貯蔵庫	ボール貯蔵庫
缶搬送機	缶搬送機
搬送回数カウンター	缶密閉機
缶残量センサ	ボール供給機
バーコード読取機	シングル缶
ボールセンサ	サンプル缶
ふたセンサ	缶
缶密閉機	缶仕様
制御パネル	シングル缶仕様
ボール供給機	サンプル缶仕様
受信	テニスボールタイプ
出力	ボール補充要求
送信	
缶	
テニスボール	

熟練者の経験知識を活用できるようにするには、オブジェクト抽出の経験知識から共通部分をうまく抽出する事が必要である。そこで、オントロジーはオブジェクトのもっと動的な側面等からも捉えることができれば、精度の高いオブジェクト抽出基準が生成可能になると考えられる。

また、抽出基準や手続きは膨大化、複雑化した場合に手順を見落とされる可能性がある。そこで、手続きに必要なタスクを精緻化する必要がある。

6. おわりに

オブジェクト抽出に使用される経験知識をオントロジーで表わすことにより誰でも良いオブジェクト抽出することを可能にする事を示した。しかし、より高いレベルの抽出結果を

得るために多様なシステムの分野に関する抽出基準を作成し、相互に比較検討することにより、抽出基準の精度向上を図る。

参考文献

- [1] 佐伯元司 他：自然言語仕様からモジュール構造を抽出する手法について、ソフトウェア工学、57-2(1987)
- [2] 滝沢陽三 他：オブジェクト指向に基づく要求記述からの形式的仕様の導出手法、ソフトウェア工学、94-8(1993)
- [3] 原田実 他：自然語要求仕様からオブジェクト指向設計図を自動生成するシステム CAMEO、情報処理学会論文誌、Vol.38 No.10(1997)
- [4] 榎木浩 他：オントロジーによる通信サービス要求定義手法、電子情報通信学会論文誌、Vol.J80-B-I No.3(1997)
- [5] 中谷多哉子 他編：ソフトウェア・パターン、pp.44-54、共立出版(2001)
- [6] J.ランポー 他著、羽生田栄一 監訳：オブジェクト指向方法論 OMT、トッパン(1992)
- [7] S. シュレイアー, S.J. メラー：続オブジェクト指向システム分析 - オブジェクトライフサイクル -、近代科学社 (1995)
- [8] 溝口理一郎：オントロジーと知識処理、情報処理学会会誌(2000)
- [9] Michael Gruninger, Jintae Lee :「 ONTOLOGY APPLICATIONS and DESIGN 」、COMMUNICATIONS OF THE ACM、Vol.45 No.2(2002)
- [10] 來村徳信 他：故障オントロジー - 概念抽出とその組織化 -、人工知能学会誌、Vol.14 No.5(1999)
- [11] 西田豊明 他編：工学知識のマネジメント、pp.155-179、朝倉書店(1998)
- [12] 武内惇 他：制限自然言語による要求記述の有効性評価、ソフトウェア工学、98-3(1994)