

# 公理としてのドメインモデルに基づく意味の規定について

## On Specifying meaning based on a domain model used as the indefinable basis

山田隆弘<sup>\*1</sup>  
Takahiro Yamada

<sup>\*1</sup> 宇宙航空研究開発機構宇宙科学研究所  
Japan Aerospace Exploration Agency/Institute of Space and Astronautical Science

None of the existing semantic theories can define the meaning of language expressions satisfactorily. In the first place, it is not possible to define the meaning of all language expressions with language expressions. Therefore, we must admit that the meaning of some words and/or expressions is indefinable and that the meaning of the other words/expressions should be defined using these indefinable words/expressions. This paper proposes a method of specifying the meaning of words/expressions using a domain model as the basis of such indefinable words/expressions, and shows that such specifications can facilitate automatic processing of natural language text.

### 1. はじめに

言語の意味に関する研究は色々と存在するが[Cann 09, Kearns 11 など]、言語表現の意味を必要十分に定義する方法として広く認められているものは存在しない。例えば、形式意味論においては、文の意味を論理式として表現する。「 $x$  は  $P$  である」という文の意味は、 $P(x)$ という述語論理式で表されるが、述語  $P$  の意味を陽に定義する方法は形式意味論では与えられていない。

そもそも、あらゆる言語表現の意味を言語表現によって定義するのは原理的に不可能であるので、基本的な語あるいは表現の意味は、定義しなくても理解できているとし(すなわち、それらの語や表現を数学における公理とみなし)、そのような語や表現を用いて他の語や表現の意味を定義する必要がある。そのような方法として知られているものに Natural Semantic Metalanguage (NSM)と言われているものがある[Wierzbicka 96]。しかし、この方法にも様々な問題点が存在する。

言語の意味を機械により自動的に処理する場合、言語表現の意味が形式的に表されていると意味の論理的な処理が可能になり、処理の精度が上がるのが期待できる。本稿では、言語表現の論理的な処理を可能にするような言語表現の意味の規定方法を提案する。この方法では、基本的な語あるいは表現の意味は定義しなくても理解できているとする(すなわち、公理とみなす)という立場に立つが、NSMとは異なる方法により公理を設定する。この方法では、ある特定の分野あるいは領域ごとに基本概念をドメインモデルとして定義し、このドメインモデルで使われている概念の意味は、定義なしに了解されているとする。個々の語や表現については、ドメインモデルで定義された概念と意味規定用のメタ言語(これも意味は了解されている、すなわち、公理の一部であると考え)を用いて規定される。なお、本稿の内容は、著者の以前の研究[山田 15]を一般化しつつ理論的に精密にしたものである。

以下、第2節において従来の意味論における問題点を指摘する。第3節ではドメインモデルを定義し、第4節ではドメインモデルに基づいて語や表現の意味を規定する方法を示す。第5節で本方式の特徴について説明し、第6節で将来の課題を示す。

### 2. 従来の意味論の問題点

#### 2.1 形式意味論

言語の意味に関する理論として最もよく知られているものは形式意味論である[Cann 09]。形式意味論においては、文の意味を論理式として表現する。例えば、「太郎は丸々電気に勤めている」という文の意味は、

(1) 勤める(太郎, 丸々電気)

という論理式で表される。ここで、「勤める」は述語であり、「太郎」と「丸々電気」は項である。項は、ある世界に存在する対象物を指し示すものとして解釈され得るが、述語「勤める」の意味を陽に定義する方法は、形式意味論では与えられていない。

形式意味論における中心的な存在であるモデル理論の意味論においては、述語「勤める」の意味は、勤める( $x, y$ )が真となるような全ての( $x, y$ )の集合によって与えられるとされる。しかし、この理論においては、このような集合はすでに与えられていることになっていて、この集合の構成規則、すなわち、勤める( $x, y$ )が真となるような( $x, y$ )の決め方は不明である。

ただし、この集合の持つ特徴を論理式によって規定することは、ある程度可能である。例えば、勤める( $x, y$ )が真となるのであれば(すなわち、 $x$  が  $y$  に勤めているのであれば)、 $x$  は人間であり、 $y$  は組織である必要があるが、これは以下のように表すことができる。

(2) 勤める( $x, y$ )  $\rightarrow$  人間( $x$ )  $\vee$  組織( $y$ )

また、 $x$  が  $y$  に勤めているのであれば、 $x$  は  $y$  から給料をもらっているので、

(3) 勤める( $x, y$ )  $\rightarrow$  受け取る( $x, 給料, y$ )

という規定を設けることもできる。

しかし、このように述語の特徴の規定を行っても、それで述語の意味が必要十分に規定されているかどうかは簡単には決められない。また、新たに使用された人間( $x$ )、組織( $y$ )などの述語の定義は不明であるし、給料という概念の定義も不明である。これらの概念が定義できたとしても、それらの定義で使用されている概念を定義する必要があり、どこまで行けば終わりになるのか定かでない。

従って、形式意味論は、論理的な処理が行えるという点で魅力的であるが、必ずしも意味の定義が行えるわけではない。

## 2.2 Natural Semantic Matalanguage (NSM)

Natural Semantic Matalanguage (NSM) [Wierzbicka 96]では、基本的な語 (Semantic Primitive と呼ばれる) を幾つか規定し、Semantic Primitive の意味は定義不可能であるとする。Semantic Primitive の集合は、しばしば改定されているが、だいたい 60 前後の語が Semantic Primitive として規定されている。Semantic Primitive 以外の語の意味は、Semantic Primitive のみを使って定義される。この方法によると語の意味を陽に定義できることになる。

しかし、Semantic Primitive を用いた意味の定義は、十分に満足できるような意味定義になっているかどうか分からないという批判は多い[Riemer 10]。例えば、ある語を定義するとき、その語が使われるあらゆる文脈において通用するような定義を確立しようとする、定義が非常に長くなる。長い定義は分かりにくく不便であり、その定義をどのように活用できるのかも不明である。

一つの例として、NSM における日本語の「甘え」の定義 [Wierzbicka 96]を以下に示す(原文は英語)。

X は以下のようなことを考える  
 Y が私のことを考える時は Y は何か良いと感じる  
 Y は私のために良いことをしたい  
 Y は私のために良いことができる  
 私が Y の近くにいる時は悪いことは私には起こり得ない  
 このために私は何もする必要がない  
 私は Y の近くにいたい  
 このために X は何か良いと感じる

このような定義が「甘え」の意味を必要十分に表しているかどうかは簡単には判定できないし、このような定義をどのように活用できるのかも不明である。

従って、NSM は、意味の定義を行っているという点で魅力的であるが、必ずしも論理的な処理には向いていない。

## 3. ドメインモデル

### 3.1 ドメインモデルの概念

本稿で提案する方法では、ある特定の分野あるいは領域ごとに基本概念をドメインモデルとして定義し、このドメインモデルで使われている概念は、定義なしに(すなわち、公理として)意味が了解されているものとする。個々の語や表現については、ドメインモデルで規定された概念と意味規定用のメタ言語(これも意味は了解されている、すなわち、公理の一部である)のみを用いて規定される。

ドメインモデルで定義されている概念は、特定の文化に基づくようなものを除いて個別言語には依存せずに定義するものとする。個別言語における語や表現は、第 4 節の方法によって言語毎にドメインモデルの概念と対応付けられる。「ドメインモデルはどのような言語で定義すべきか」という問題もあるが、特定の文化に依存しない概念であれば、任意の言語で定義可能である。以下では、日本語でドメインモデルを定義する。

以下では、ドメインモデルの定義とその問題点を示す。意味規定用のメタ言語については、第 4 節で述べる。

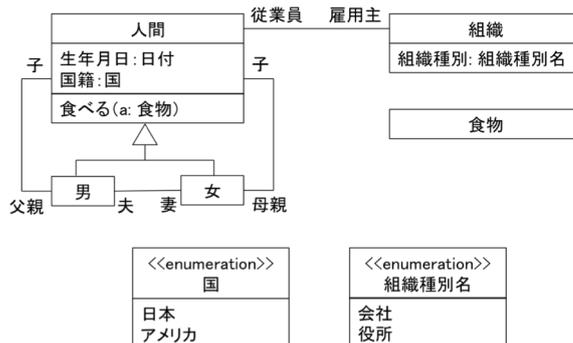


図1 UML のクラスモデルとして表したドメインモデルの例

### 3.2 ドメインモデルの定義

ここで使用するドメインモデルは、オブジェクト指向モデル [Blaha 05]に基づいているものであり、各々のドメイン毎に以下の概念を定義する。

- クラス(共通の特徴を有するオブジェクト(個物)の集合)
- アトリビュート(オブジェクトの性質を表現するためのパラメータ)
- アトリビュート値(アトリビュートが取り得る値)
- アソシエーション(同一あるいは異なるクラスに属するオブジェクト間の関係)
- オペレーション(オブジェクトが行う行為)
- オペレーションのパラメータ(行為の内容を特定するためのパラメータ)
- クラス間の包含関係(あるクラスが別のクラスを含むという関係)

上に示した概念は、Unified Modeling Language (UML) [Booch 05]のクラス図で定義できるものに対応している。また、記述論理[Baader 10]における T-Box として定義できるものにも相当する。

非常に簡単なドメインモデルの例を UML のクラス図として表したものを図1に示す。この図では、以下の概念が定義されている。

- (クラス)人間、男、女、組織、食物
- (アトリビュート)生年月日、国籍、組織種別
- (アトリビュート値)日付、日本、アメリカ、会社、役所
- (アソシエーション)親、子、夫、妻、従業員、雇用主
- (オペレーション)食べる
- (オペレーションのパラメータ)('食べる'オペレーションの)食物
- (クラス間の包含関係)人間は男と女を含む

図1に示したドメインモデルは、人間社会のごく一部を表したものであるが、人間が組織(個人経営の商店等も含む)に雇用されるという状況は、世界のかなり多くの文化で見られることであり、このモデルは、世界の多くの言語の意味規定に使用することができると思われる。

### 3.3 ドメインモデル定義上の問題点

ドメインモデルを定義するときの問題点の一つは、一つの分野において複数の異なるドメインモデルを作ることができることである。例えば、図1のドメインモデルでは、国籍は人間のアトリビュートとなっているが、国というクラスを設定し、国籍を人間クラスと国クラスの間のアソシエーションとして定義することもできる。しかし、国クラスのオブジェクトに対してアトリビュートを定義する

必要がない場合は、国クラスを定義する必要はない。ドメインモデルの用途(すなわち、どのような範囲の言語表現の意味規定にドメインモデルが使用されるか)に応じてドメインモデルを一意に決めるための基準を作成することは、将来の課題である。

また、一つの分野の様々な用途に適用できるような標準的なドメインモデルを定義し、用途毎に標準モデルを変形させて使用することが可能になれば、用途毎に別々にドメインモデルを作成する手間が省けるし、様々な用途に共通に使用できるような意味規定も作成することができるであろう。そのようなドメインモデルの構築方法の検討も、今後の課題である。

また、4.2 節で述べるようにドメインモデルで定義されている概念より論理的に派生できる概念は、ドメインモデルに含める必要がない。したがって、ドメインモデルは、そのドメインにおける意味規定に必要な概念を全て含みつつ最小なものを定義する必要がある。

## 4. ドメインモデルに基づく意味の規定

### 4.1 意味規定の方法

特定の分野や領域において使用される特定の言語の語や表現の意味は、ドメインモデルで定義されている概念を使用して規定される。この規定の方法には、4.2 節で述べるように幾つかの種類があるが、基本的な考え方は、特定言語の表現のパターンとドメインモデルに基づいた意味の規定との間の1対1の対応を付けることである。ところで、著者は[山田 17]において自然言語を自動処理できるための条件として、特定言語における言語表現のパターンと意味構造との間の1対1の対応表が必要であると述べたが、本稿で提案する方法は、その対応表を作るためのものである。

ここで言語表現のパターンについて説明する。著者が[山田 16]で論じたように、自然言語は曖昧性(一つの語あるいは表現が複数の意味で使用されること)を有しているが、その理由は自然言語が情報伝達の経済性(短い発話で多くの情報を伝達すること)を追求しているからである。この曖昧性は、語を周囲の語から切り離して考えるのではなく、複数の語のつながりによって解消する場合がある。このような場合は、語のつながりのパターンに対して一つの意味規定を対応させる。

一つの例を示す。「太郎は丸々電気で働いている」は、「太郎は丸々電気(組織)の従業員である」と解釈できる。しかし、「太郎は京都で働いている」は、「太郎は京都(場所)で何らかの仕事をしている」と解釈すべきである。同じ「働く」という語の意味を規定する場合、「組織で働く」のか「場所で働く」のかを区別しなければならない。このような場合における語のつながり方ここではパターンと呼んでいる。上記の場合、「X(人間)は Y(組織)で働く」というパターンの意味は、「X(人間)は Y(組織)の従業員である」と規定されることになる。この意味規定では、図1に示したドメインモデルの概念だけを使って規定されていることに注意されたい。また、ここで使用されている「XはYのZである」という構文は、意味規定用のメタ言語に属している。このように本方法では、公理のみを使用して意味の規定がなされる。

### 4.2 意味規定の実例

ドメインモデルに基づく意味規定の幾つかの実例を表1に示す。この表の「言語表現」の欄は、意味規定の対象となる日本語あるいは英語の語、句、文、あるいは、それらのパターンである。また、「ドメインモデルに基づく意味規定」の欄は、図1のドメインモデルで定義されている概念と意味規定用のメタ言語を用いて表した意味規定である。意味規定用のメタ言語の厳密な文法は、

本稿では示していないが、これを規定することは将来の重要な課題である。

意味規定の最初の種別は、特定言語の一つの語あるいは句の意味がドメインモデルで定義されている何らかの概念、あるいは、ドメインモデルで定義されている概念の下位概念に対応する場合である。このような場合は、語あるいは句に対応する概念を意味規定とすれば良い。表1の最初から4番目の例がこれに相当する。

二つ目の種別は、特定言語の一つの語あるいは句の意味がドメインモデルで定義されている概念より論理演算によって派生される概念に対応する場合である。この場合は、ドメインモデルの概念に対する論理式を意味規定とすれば良い。表1の5番目から7番目の例がこれに相当する。

三つ目の種別は、特定言語の表現パターンの意味がドメインモデルで定義されている概念に相当する個物(オブジェクト)の存在や属性に対応する場合である。この場合は、著者が以前[山田 14]において論じた状況(state)の意味構造の記述方法を利用して意味規定を作成できる。表1の8番目から10番目の例がこれに相当する。

四つ目の種別は、特定言語の表現パターンの意味がドメインモデルで定義されている概念に相当する個物(オブジェクト)に関する事象に対応する場合である。これは著者が以前[山田 14]において論じた事象(event)の意味構造の記述方法を利用して意味規定を作成できる。表1の11番目から13番目の例がこれに相当する。

表1 ドメインモデルに基づいた意味規定の例

意味規定種別	言語表現	ドメインモデルに基づく意味規定
ドメインモデルの概念に対応	人間	人間
	人	人間
	会社	組織種別が会社である組織
	日本人	国籍が日本である人間
ドメインモデルから派生される概念に対応	親	父親または母親
	X(人間)の兄	Xと父親か母親の少なくとも一方を共有し、生年月日がXよりも早い男
	Brother of X	Xと父親か母親の少なくとも一方を共有する男
ドメインモデルの概念に相当する個物の状況に対応	X(人間)は日本人である	Xの国籍は日本である
	X(人間)はY(組織)に勤めている	YはXの雇用主であり、XはYの従業員である
	X(男)とY(女)は結婚している	YはXの妻であり、XはYの夫である
ドメインモデルの概念に相当する個物に関する事象に対応	X(人間)が生まれる	人間クラスのオブジェクトXが生成される
	X(人間)はY(組織)に就職する	YはXの雇用主に、XはYの従業員になる
	X(人間)はY(食物)を食べる	XはYをパラメータとして食べるオペレーションを実行する

## 5. 本方式の特徴

ここで提案した方式の特徴は、あるドメインにおけるあらゆる言語表現の意味をドメインモデルで定義された概念に還元して表現することができるので、そのドメインにおけるあらゆる言語表現の間の関係が厳密に規定できることである。これにより該当ドメインの文章に基づいた自動的な推論が可能になる。例えば、「太郎は花子と結婚している」という文と「花子は四越百貨店に勤めている」という文が与えられている時に、「太郎の妻の勤務先は？」という問いに対して「四越百貨店」と回答することができる。また、「太郎は花子と結婚している」という文と「花子の夫は次郎である」という文とが矛盾していることも判定できる。従って、本方式は、自動的に言語処理を行うための論理的な基盤を与えることになる。ただし、本方式によってどの範囲の推論が可能になるのか等については、詳しい検討が今後必要になるであろう。

また、自然言語の文をドメインモデルに整合するように作文すれば、上で述べた自動処理が行いやすくなる。従って、自動処理したい文書を作成するときは、該当分野のドメインモデルとそれに従って作成された意味規定(表1のようなもの)を参照して作成すべきである。また、ドメインモデルに整合している文章を書くためのツールなども整備すべきである。世の中には、自動処理できると便利であるような文書は、極めて多数存在しているので、そのような文書を作成するために用いるというのも本方式の重要な用途である。

この方式には、問題点も存在する。まず、この方式を実用的な文章に当てはめようとすると、ドメインモデルが大きくなるという問題がある。3.3 節で述べたように、ドメインモデルで定義されている概念より論理的に派生できる概念は、ドメインモデルに含められないようにする必要があるが、アトリビュートやオペレーションは論理的に定義できないものが多い。その場合は、分かりやすいドメインモデルを構築し、ドメインモデルを使いやすくするための工夫が必要である。

本稿では、自然言語の文における述語と項に関する意味規定のみを考えてきたが、自然言語の文の意味要素には他にも様々な要素が存在する。例えば、時制や法(・すべきなどの表現)の表現方法や複数の文の間の関係の表現方法については、おおまかな意味表現の方法を[山田 14]にて提案したが、まだ自動処理できるほど厳密なものにはなっていない。このような様々な意味要素の厳密な表現方法の開発は、将来の課題である。

本稿では、曖昧性解消のための一つの方法を提示したが、これであらゆる曖昧性が解消できるわけではない。複数の文にまたがる曖昧性の解消(例えば、代名詞の先行詞の決定など)も必要である。本稿の方式は、談話(複数の文)の意味の規定に発展させる必要があり、これも将来の課題である。

## 6. 将来の課題

この節では、本研究に関する将来の課題をまとめることにする。

まず、ドメインモデル構築に関する課題としては、それぞれの分野において様々な用途に適用できるような標準的なドメインモデルの構築方法を検討することがある。また、各々の分野における必要な概念を網羅しながら最小なドメインモデルを構築する方法、および、分かりやすく使いやすいうドメインモデルを構築する方法の検討も必要である。実用に耐えるドメインモデルの実例を構築することも重要な課題である。

意味規程の側の問題としては、意味規定用のメタ言語(これは公理の一部となる)の厳密な定義を開発する必要がある。本稿では省略した時制や法のような意味要素の表現方法や談話の意味を規定する方法も必要となる。

また、本方式の応用に関しては、本方法がどのような用途(質問応答など)にどれだけ役に立つかの評価が必要である。また、本方式を有効に利用するためには、ドメインモデルに整合した文書を作成する必要があり、それを行うためのツールの開発も必要となる。

## 7. おわりに

本稿では、言語の意味定義の一つの方法としてドメインモデルを公理(定義なしに意味が了解されているもの)として使用する方式を提案した。この方法は、自然言語を自動的に処理するための論理的な基盤を与えるものになると思われる。この方法は、著者の今までの断片的な研究([山田 14]、[山田 15]、[山田 17]等)を一つの体系としてまとめたものであるが、第 6 節で述べたように将来の課題も多い。また、本稿で提示した方法は、言語による情報伝達機能に関するさらに大きな理論の一部にもなる予定である。今後は、本稿の方法の詳細を詰めるとともに、さらに大きな理論も展開させたいと思っている。

## 参考文献

- [山田 14] 山田隆弘: オブジェクト指向に基づく意味構造の記述について, 2014 年度人工知能学会全国大会, 214-OS-08a-4 (2014).
- [山田 15] 山田隆弘: Linked Data における述語の意味を規定する方法, 人工知能学会研究会資料, SIG-SWO-037-05 (2015).
- [山田 16] 山田隆弘: 意味の理解とは何であるのか - 通信理論からのアプローチ -, 2016 年度人工知能学会全国大会, 303-OS-04a-4 (2016).
- [山田 17] 山田隆弘: 自動処理できる言語とはどのようなものか, 言語処理学会第 23 回年次大会, D1-5, 2017.
- [Baader 10] Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., Patel-Schneider, P. F.: *The Description Logic Handbook: Theory, Implementation and Applications*, Second Edition, Cambridge University Press, 2010.
- [Blaha 05] Blaha, M., Rumbaugh, J.: *Object-Oriented Modeling and Design with UML*, Second Edition, Prentice Hall, 2005.
- [Booch 05] Booch, G., Rumbaugh, J., Jacobson, I.: *The Unified Modeling Language User Guide*, Second Edition, Addison-Wesley, 2005.
- [Cann 09] Cann, R., Kempson, R., Gregoromichelaki, E.: *Semantics*, Cambridge University Press (2009).
- [Kearns 11] Kearns, K.: *Semantics*, Second Edition, Palgrave Macmillan (2011).
- [Riemer 10] Riemer, N.: *Introducing Semantics*, Cambridge University Press (2010).
- [Wierzbicka 96] Wierzbicka, A.: *Semantics, Primes and Universals*, Oxford University Press (1996).