

形式オントロジーと語彙意味論を用いた自然言語推論の試み

理学専攻・情報科学コース 中村 絢子

1 はじめに

自然言語の含意関係認識では、推論の前提や結論の意味内容だけでなく、前提には明示されていない世界知識を必要とする場合があることはよく知られている。

- (1)

前提： ラッセル車が運行される。
結論： 除雪作業を行う列車が運行される。

(1) は、「ラッセル車は除雪作業を行う列車である」という知識を使って前提から結論が導かれる例である。

また、照応現象の一つであるいわゆる bridging (Clark [3]) においても同様の世界知識が必要になる。

- (2) 北斗星に乗って夕食を食べようと思った。
しかし、食堂車はまだ開いていなかった。

(2) において第 2 文の「食堂車」を「北斗星の食堂車」と解釈するためには、「北斗星には食堂車がある」という知識を用いる必要がある。

このような知識の形式的記述には、オントロジーを用いることが有効である。溝口ら [7] によると、オントロジーとは、共通語彙 (概念) を提供する体系的な辞書であり、概念と概念間の関係 (意味リンク) を用いて記述される。概念間の関係を明示的に記述することによって、そのバックグラウンドにある暗黙的な情報を明示することが可能となる。人工言語ベースの形式的なオントロジーは、高い形式性をもち、従来の上位・下位概念に基づく概念関係の記述を超えたより精緻な知識の構造的記述が可能になるという利点がある。また、従来の形式意味論は、文の意味論の枠組みとして発展し、さらに談話表示理論 (DRT, Kamp [4]) などの出現によって、80 年代以降は談話の意味論 (discourse semantics) としても大きな発展を遂げているが、これを語彙意味論 (lexical semantics) と接合する試みはいまだ十分になされていない (cf. Asher and Pustejovsky [1])。

本論文は、溝口ら [8][7] の形式的なオントロジーに基づいて、動詞に加えて名詞によって表現されるより複雑な概念間の関係についても記述することを目的とし、自然言語推論に必要な世界知識を記述する枠組みとはどのようなものになるのかを考察する。さらに、DTS の型理論的な枠組みに世界知識・語彙知識に基づく推論を扱うメカニズムを結合することで、形式意味論と語彙意味論とを組み合わせたい研究への貢献を目指すものである。

2 オントロジーの構築

ここでは、is-a リンク、part-of リンク、attribute-of リンクの 3 種類の意味リンクを取り上げる。列車は、「気動車」や「電車」等の異なる概念を含んでいる。これらを分けて記述する場合は、列車の集合の部分集合として気動車の集合や電車の集合があることを明記する必要がある。これを is-a リンクを用いて記述する。is-a リンクの元を上位概念 (スーパークラス)、is-a リンクの先を下位概念 (サブクラス) と呼ぶ。is-a リンク

を使用して種類を書き分ける場合は、部分集合同士が disjoint になるようにする必要がある。

各概念がもっている性質は、part-of リンクと attribute-of リンクを用いて記述される。「列車は、扉、動力、ブレーキを部分としてもち、重さ・長さを属性としてもつ」という知識は階層的に表される。part-of リンク (p/o) は、「扉」や「ブレーキ」など、「列車」の空間的な部分を構成する概念に対して用いられる。一方、重さ・長さ等の物理量や色は、attribute-of リンク (a/o) を用いて記述される。

「列車」「気動車」「電車」「人間」など、コンテキストや他の概念に依存せず、それ自身の性質に基づいて規定される概念は基本概念と呼ばれる。一方、コンテキストに依存して決定される役割を表す概念は、ロール概念と呼ばれる。例えば、「乗客」「操縦者」は、ロール概念であり、「電車」という概念が定めるコンテキストの中での役割を表す。ロール概念が規定する役割を担う基本概念は、クラス制約によって指定される。

3 世界知識と語彙知識の変換

オントロジーによって記述された世界知識と語彙知識が、どのように推論・照応の計算に使用できる形に変換されるか考察する。

3.1 依存型意味論 (DTS) とオントロジーの変換

推論を行うために、Bekki [2] で提案された依存型意味論 (Dependent Type Semantics, DTS) を用いる。DTS は、依存型に基づく自然言語の意味論であり、含意関係の導出と照応の解消を「証明の構成」として統一的に扱える点に特徴がある。DTS が自然言語の意味論として優れている点は、以下の 4 点である。

1. 動的な束縛を表現することができる。
2. 証明論的に計算することができる。
3. 語彙化文法の意味部門として利用できる。
4. 到達可能性の条件を表現することができる。

依存型理論では、 Π 型、 Σ 型が用いられる。 Π 型は全称量化に対応し、 Σ 型は存在量化に対応する。

	Π 型	Σ 型
標準的な記法	$(\Pi x : A) B(x)$	$(\Sigma x : A) B(x)$
DTS の記法	$(x : A) \rightarrow B(x)$	$\left[\begin{array}{l} x:A \\ B(x) \end{array} \right]$
$x \notin fv(B)$ の場合	$A \rightarrow B$	$\left[\begin{array}{l} A \\ B \end{array} \right]$

この表記を用いて、前節で導入したオントロジーの 3 種類の意味リンクは、それぞれ以下のように DTS に翻訳される。


1. is-a リンク

ラッセル車は列車である

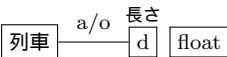
ラッセル車	is-a	列車
-------	------	----

 $(x : \text{Entity}) \rightarrow (\text{ラッセル車}(x) \rightarrow \text{列車}(x))$

2. part-of リンク

ラッセル車は除雪装置をもつ 
 $(x : \text{Entity}) \rightarrow (\text{ラッセル車}(x) \rightarrow \left[\begin{array}{l} y : \text{Entity} \\ \text{除雪装置}(y, x) \end{array} \right])$

3. attribute-of リンク

列車は float 型の長さ d をもつ 
 $(x : \text{Entity}) \rightarrow (\text{列車}(x) \rightarrow \left[\begin{array}{l} d : \text{float} \\ \text{length}(x) = d \end{array} \right])$

3.2 語彙知識と DTS の公理への変換

自然言語の推論の導出では、世界知識だけでなく、個々の語彙の意味に関する知識（語彙知識）も重要な役割を果たす。この節では、形容詞的述語の語彙知識に注目する。*dirty* と *clean* のような反義語のペアは、部分と全体の関係に依存した推論を可能にすることが知られている (Yoon[6])。

- (3) a. Are the toys dirty?
 b. Are the toys clean?

例えば、(3a) では、*the toys* が指すおもちゃの集まりのうち、その一部が汚れていれば *yes* と答えることができるが、(3b) の場合、おもちゃの全体がきれいではないと *yes* とは言いにくい。*dirty* を部分述語 (partial predicate)、*clean* を全体述語 (total predicate) と呼ぶ [6]。

部分述語と全体述語にかかわる語彙知識はそれぞれ次のような DTS の公理として記述される。

- (4) 全体述語 (正常だ)
 $(x : E) \rightarrow (y : E) \rightarrow \text{正常 } x \rightarrow y \preceq x \rightarrow \text{正常 } y$
- (5) 部分述語 (故障する)
 $(x : E) \rightarrow (y : E) \rightarrow \text{故障 } x \rightarrow x \preceq y \rightarrow \text{故障 } y$

全体述語は、複合物全体について成り立つならその部分についても成り立つ。一方、部分述語は部分について成り立つならそれを含む全体についても成り立つ。

4 オントロジーを用いた DTS による推論

世界知識と語彙知識を用いた推論がどのように行われるかを考察する。次の推論を取り上げる。

- (6) $\frac{\text{前提 : 列車は正常である}}{\text{結論 : 車輪は正常である}}$

前提と結論はそれぞれ (7) と (8) に示される DTS の意味表示に対応する¹。ここでは、前提文の「列車」が指示する対象を個体の型 E をもつ変数 a によって表している。

$$(7) \left[\begin{array}{l} \text{列車 } a \\ \text{正常 } a \end{array} \right]$$

$$(8) (x : E) \rightarrow \left[\begin{array}{l} \text{車輪 } x \\ x \preceq a \end{array} \right] \rightarrow \text{正常 } x$$

¹結論文の「車輪」は、中村ら [10] で論じた bridging により、「列車の車輪」と解釈する必要がある。単純化のため、ここでは照応・前提がすでに解消された後の意味表示を想定している。

「正常である」は全体述語であるから、(4) の公理が使用可能である (その証明項を f とする)。この他に、 $a : E$ と前提に対応する (7) の型をもつ証明項 (これを p とする) の存在を仮定し、結論 (8) の証明項は以下になる。

$$\lambda q \lambda x. f a x (\pi_2 p) (\pi_2 q) : (x : E) \rightarrow \left[\begin{array}{l} \text{車輪 } x \\ x \preceq a \end{array} \right] \rightarrow \text{正常 } x$$

次に部分述語を含む推論の例を検討する。

- (9) $\frac{\text{前提 : 車輪が故障している}}{\text{結論 : 列車が故障している}}$

前提と結論はそれぞれ、(10) と (11) の意味表示に翻訳される。

$$(10) \left[\begin{array}{l} u : \left[\begin{array}{l} x : E \\ \text{車輪 } x \end{array} \right] \\ \text{故障 } \pi_1 u \end{array} \right]$$

$$(11) \left[\begin{array}{l} u : \left[\begin{array}{l} x : E \\ \text{列車 } x \end{array} \right] \\ \text{故障 } \pi_1 u \end{array} \right]$$

部分述語である「故障している」には、(5) の公理が適用可能である。また、「列車は車輪を一部としてもつ」というオントロジーの part-of リンクによって記述された知識から次の公理が得られる。

$$(12) (x : E) \rightarrow \text{車輪 } x \rightarrow \left[\begin{array}{l} y : E \\ \text{列車 } y \\ x \preceq y \end{array} \right]$$

前提 (10) の証明項の存在を仮定すると、結論の型 (11) の証明項を構成することができ、(9) の推論が導出可能である。

参考文献

- [1] N. Asher and J. Pustejovsky. The metaphysics of words in context. *Journal of Logic, Language and Information*, 2000.
- [2] Daisuke Bekki. Representing anaphora with dependent types. In *Logical Aspects of Computational Linguistics*, pp. 14–29. Springer, 2014.
- [3] Herbert H. Clark. Bridging. In *Theoretical issues in natural language processing*, pp. 169–174, 1975.
- [4] Hans Kamp and Uwe Reyle. *From Discourse to Logic; Introduction to the Modeltheoretic Semantics of natural language*. Reidel, 1993.
- [5] Ayako Nakamura, Koji Mineshima, and Daisuke Bekki. Towards modeling natural language inferences with part-whole relations using formal ontology and lexical semantics. 2015. 6 pages.
- [6] Youngeun Yoon. Total and partial predicates and the weak and strong interpretations. *Natural language semantics*, Vol. 4, No. 3, pp. 217–236, 1996.
- [7] 溝口理一郎. オントロジー工学. オーム社, 2005.
- [8] 溝口理一郎, 古崎晃司, 笹島宗彦, 来村徳信. オントロジー構築入門. オーム社, 2006.
- [9] 中村絢子, 峯島宏次, 戸次大介. オントロジーを用いた型理論的な語彙意味論の試み. 4 pages.
- [10] 中村絢子, 峯島宏次, 戸次大介. オントロジーを用いた自然言語の推論に向けて. 言語処理学会第 21 回年次大会発表論文集, pp. 309–312, 2015.