

# 依存型意味論を用いた日本語含意関係コーパスの分析に向けて

中村 絢子<sup>1</sup>

金子 貴美<sup>2,3</sup>

戸次 大介<sup>2,3,4</sup>

<sup>1</sup> お茶の水女子大学 理学部 情報科学科

<sup>2</sup> お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻 情報科学コース

<sup>3</sup> 国立情報学研究所

<sup>4</sup> 独立行政法人科学技術振興機構, CREST

{g1020525, kaneko.kimi, bekki}@is.ocha.ac.jp

## 1 はじめに

2つの文が含意関係を持つとは一方の文が真であるときに他方の文も真である<sup>1</sup>ということを示す。自然言語処理において含意関係を自動処理する必要性が高まっている。文間の含意関係の有無を判定するシステムに機械学習させるための正解データを構築する方法論として金子ら [1] のものがある。以下に金子ら [1] で挙げられた例を示す。

- (1) a. プレイヤーはパックマンを直接操作する。
- b. プレイヤーは直接パックマンを操作できる。

(1a) と (1b) はほぼ同じ内容を示しているが、語順や接尾語などが異なっている。

また、一方の文が他方の文を含意している例を示す。以下の例では (2a) は (2b) を含意している。

- (2) a. バスケットボールの欧州選手権は、1935年に第1回大会が開催された伝統ある大会である。
- b. バスケットボールの欧州選手権は、1935年に第1回大会が開催された。

現時点では、正解データの構築手法は人間の直観に基づいて策定されており、計算機に自動認識させる際に金子ら [1] の手法で提案された分類が言語学的に妥当であるかどうかの分析がなされていない。

そこで、本研究では金子ら [1] の正解データの例文に Bekki[2] で提案された依存型意味論によって意味表示を与え正解データの構築手法を分析する。また、推論

<sup>1</sup>自然言語処理においては“含意関係”という用語はもう少し広い意味で用いられている。

の際に必要な知識を依存型理論を用いたオントロジーで記述する。

## 2 関連研究

本節では含意関係認識の機械学習のための正解データを構築する手法の提案として金子ら [1] の研究および Bekki[2] で提案された DTS について述べる。

### 2.1 金子ら [1]

金子ら [1] は、正解データの構築を通して、1つの文を複数回書き換えて目的の文を生成する過程を提示している。また、1度書き換えるごとに書き換える前の文と書き換えた後の文との間の関係に基本文関係をつけている。基本文関係とは、2文間に成立するプリミティブな含意関係である。任意の含意関係は任意個の基本文関係で表される。以下に金子ら [1] で提示されたカテゴリーの例を示す。

カテゴリーラベル	例
同義・類義 (単語の書き換え)	「法廷で → 法廷において」
含意・前提 (フレーズの書き換え)	「ノーベル文学賞を受賞した → 作者である」
時間の推論	「1931年 → 1930年代」

表 1: 基本文関係の例

また、文の書き換えの例を示す。⟨t2⟩ は ⟨t1⟩ の書き換えである。

- ⟨t1⟩ 新卒一括採用では年功序列を基礎として、年長者が上司、若年者が部下となり、指揮・命令系統に不備がでない。⟨\t1⟩
- ⟨t2⟩ 新卒一括採用では年功序列を基礎として、年長者が上司となる。⟨\t2⟩

これらは人手で書き換え・分類されており、書き換える道筋が複数存在することが考えられる。その場合に、書き換える結果が一致するかどうかの分析がなされていない。また、文を書き換える際に統語的なレベルで書き換える場合と意味的なレベルで書き換える場合が混在していることも正解データの妥当性を分析することを困難にしていると考えられる。

## 2.2 Bekki[2]

Bekki[2] で提案された DTS は、依存型理論を用いた自然言語の意味論である。

依存型理論は、「項に依存した型」を表現することが可能な体系である。例えば、型 Month の項 Jan, Feb, Mar... に対してそれぞれ Day(Jan), Day(Feb), Day(Mar)... のような型の集合を考えると、日付は「型が Month であるような項 m」と「型が Day(m) であるような項」の組で表される。これにより「2月30日」のような実在しない日付を排除することができる。

DTS が自然言語の意味論として優れている点は以下の4項目である。

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。

実テキストを網羅した意味論は多くなく、DTS が実テキストにおいてどの程度の範囲を網羅しているかについては検証が待たれる。

## 3 分析

これらの点を踏まえ、本研究では、金子ら [1] で提案された正解データの例文に Bekki[2] で提案された DTS による意味表示を与え、金子ら [1] の正解データの構築手法を分析すると共に、DTS の実テキストに対する経験的な網羅性を検証する。

以下の4文 ((3)~(6)) に DTS を用いて意味表示を与えた ((3')~(6'))。

- (3) 年功序列を基礎として、年長者が上司、若年者が部下となり、指揮・命令系統に不備が出ない。
- (4) 年功序列を基礎として、年長者が上司となる。

- (5) ルキウス・ドミティウス・アウレリアヌス (以下ルキウス) の突然の死で皇帝不在の中、マルクス・クラウディウス・タキトゥス (以下マルクス) が帝位につく。
- (6) アウレリアヌスの突然の死で皇帝不在の中、マルクス・クラウディウス・タキトゥスが即位する。

(3')

$(\lambda c)(\Sigma w:World)(\Sigma t:Time(w))$

$\Sigma u_1:((\Sigma e:Event(w)) \text{ とする } (w,e,pro, \text{年功序列, 基礎}) \wedge pres(e,t))$

$\Sigma u_2:((\Sigma e:Event(w)) \text{ となる } (w,e, \text{年長者, 上司}) \wedge pres(e,t))$

$\wedge \text{となる } (w,e, \text{若年者, 部下}) \wedge pres(e,t))$

$\Sigma u_3:(\neg(\Sigma e:Event(w)) \text{ 不備がでる } (w,e, \text{指揮・命令系統})$

$\wedge pres(e,t))$

(4')

$(\lambda c)(\Sigma w:World)(\Sigma t:Time(w))$

$\Sigma u_1:((\Sigma e:Event(w)) \text{ とする } (w,e, \text{年功序列, 基礎}) \wedge pres(e,t))$

$\Sigma u_2:((\Sigma e:Event(w)) \text{ となる } (w,e, \text{年長者, 上司})$

$\wedge pres(e,t))$

経験的に (3) は (4) を含意している。DTS では、(3') から (4') は依存型理論の推論規則 ( $\Sigma E$ ) によって導出可能である。<sup>2</sup>よって、金子ら [1] の正解データは理論的な含意関係と一致する。しかし、次の例では説明が難しい。

(5')

$(\lambda c)(\Sigma w:World)(\Sigma t:Time(w))$

$\Sigma u_1:((\Sigma e:Event(w)) \text{ 突然の死 } (w,e, \text{ルキウス}) \wedge pres(e,t))$

$\Sigma u_2:(\neg(\Sigma e:Event(w)) \text{ 存在する } (w,e, \text{皇帝}) \wedge pres(e,t))$

$\Sigma u_3:((\Sigma e:Event(w)) \text{ つく } (w,e, \text{マルクス, 帝位}) \wedge pres(e,t))$

(6')

$(\lambda c)(\Sigma w:World)(\Sigma t:Time(w))$

$\Sigma u_1:((\Sigma e:Event(w)) \text{ 突然の死 } (w,e, \text{アウレリアヌス}) \wedge pres(e,t))$

$\Sigma u_2:(\neg(\Sigma e:Event(w)) \text{ 存在する } (w,e, \text{皇帝}) \wedge pres(e,t))$

$\Sigma u_3:((\Sigma e:Event(w)) \text{ 即位する } (w,e, \text{マルクス}) \wedge pres(e,t))$

(5) と (6) は意味的にはおおよそ同じことを述べた文であるが、名前が省略されていたり、(5) の「帝位に

<sup>2</sup> $\Sigma E$  はペアの第一要素、または第二要素を取り出す操作に相当する。

つく」という言葉と同じ意味で (6) では「即位する」という言葉が使われていたりするなど意味表示の違いが見られる。このような場合、単純に依存型理論の推論規則を適用しただけでは (5') から (6') を導出することは不可能である。そこで、「帝位につく」という言葉と「即位する」という言葉が意味が近いという知識が必要になり、またその知識を記述することが重要になる。

### 3.1 オントロジー

オントロジーとは、概念間の関係を記述したものであり、本研究では Patrick[3] に基づき部分型をもつ依存型理論を用いて記述する。例えば Patrick[3] では「リンゴは色を持つ」という概念は以下のように記述される。

$$\Pi u : (\Sigma x : \text{Apple.Color}[x]) \text{HasColor}[\pi_1 u]$$

$\Pi$  は全称量子化 ( $\forall$ ) に対応する働きをする。よってこの式は全てのリンゴは何かしらの色を持つことを表している。

これを踏まえ、(5) の「帝位につく」という概念と (6) の「即位する」という概念の関係を記述することを考える。「帝位につく」という概念は「即位する」という概念より狭い概念を表している。よって「帝位につく」という概念は「即位する」という概念の下位概念と捉えることができる。部分型を持つ依存型理論で下位概念を記述する場合は以下のように記述することができる。

帝位につく  $\preceq$  即位する

部分型を推論する場合には以下の規則が必要になる。

$$\frac{\Gamma \vdash M : A \quad \Gamma \vdash A' : \text{Type} \quad A \preceq A'}{\Gamma \vdash M : A'} \text{ (Sub)}$$

また、「帝位につく」という概念と「即位する」という概念の関係を DTS を用いて記述すると以下のようになる。

$$\begin{aligned} & (\Pi w : \text{World}) \text{Actual}(w) \rightarrow \\ & (\Pi e : \text{Event}(w)) (\Pi x : \text{Entity}(w)) \\ & \text{帝位につく}(w, e, x) \rightarrow \text{即位する}(w, e, x) \end{aligned}$$

上の記述と推論規則を利用すると (5') から (6') を導出することが可能である。

## 4 考察

(3) と (4) のように一方の文が他方の文を論理的に含意するような場合は、推論規則を適用することによって一方の意味表示から他方の意味表示を導出することができる。一方 (5)、(6) のようにほぼ同じ意味だが異なる言葉に書き換えられた 2 文の場合は、一方の意味表示から他方の意味表示を導出するために推論規則以外に知識が必要である。知識を記述する際には意味論と同じ体系である依存型理論を用いることによって知識を容易に推論に用いることができると考えられる。

## 5 まとめ

本論文では、金子ら [1] が提案した含意関係認識のための正解データの一部を Bekki[2] で提案された依存型意味論で記述するによって金子ら [1] の正解データの構築手法の妥当性と依存型意味論の実テキストに対する網羅性を検証した。一方の文が他方の文を含意するような例では依存型理論の推論規則を適用することによって比較的容易に 2 つの文が含意関係を持つことを証明することができる。また、推論規則のみでは証明が困難な類義語への書き換えなどは依存型理論を用いたオントロジーを用いることによって、意味論と同じ体系で知識の記述が可能になった。金子ら [1] で提案された手法で書き換えた場合の書き換えの結果が一致するかどうかの計算は今後の課題である。また、記述した知識を実際に意味論で推論する際に利用できるようにしていく予定である。

## 参考文献

- [1] 金子 貴美, 戸次 大介, 宮尾 祐介: 「基本文関係に分解した日本語含意関係認識アノテーション」, 人工知能学会第 27 回全国大会, (2013).
- [2] Bekki, Daisuke. (to appear). “Dependent Type Semantics: An Introduction“, the 2012 edition of the LIRa yearbook: a selection of papers, University of Amsterdam.
- [3] Patrick, B., Richard, D. 2012. “A type-theoretical approach for ontologies: The case of roles“, Applied Ontology 7 pp.311-356
- [4] 兼岩 憲. 2009. “記述論理と Web オントロジー言語 “. オーム社