

## 意味を用いた前方照応の解決

4R-2

倉内貴裕<sup>†</sup>山村毅<sup>†</sup>大西昇<sup>†‡</sup><sup>†</sup>名古屋大学大学院工学研究科情報工学専攻<sup>†</sup>愛知県立大学情報科学部<sup>‡</sup>理化学研究所 バイオミメティックコントロール研究センター

## 1 はじめに

高度な自然言語処理を行うためには、文章中の照応問題の解決が必要不可欠である。この問題については、従来、主として、文法や構文上の立場から、あるいは最近では用例ベースの立場から、さまざまなアプローチが試みられているが、文の意味表現によるものは少ない[1]-[3]。本稿では、文の意味表現を与えることにより、照応形が文中に出現している対象を直接指示していない場合についての照応（推論による照応）の解決方法を提案する。

## 2 推論による照応

従来、以下の3つの制約が照応における先行詞の同定のために用いられてきた[1]。

- 文法的制約
- 意味的制約
- 談話的制約

しかし、これらの制約だけでは、文の表層に直接的に指示対象が存在する照応（直接照応）は扱えても、下の例文のように、それが文中に明示されていない場合の照応（間接照応）を扱うのは困難である。

例文：

(S1) 硫酸第一鉄の米粒大の結晶を試験管にとり、(S2) 約5立方センチメートルの蒸留水を加えて(S3) 溶かす。(S4) これにフェリシアン化カリウム溶液を1~2滴加える。

Anaphora Resolution in Japanese Demonstrative Pronoun using Semantic Representation

Takahiro Kurauchi<sup>†</sup>, Tsuyoshi Yamamura<sup>†</sup>, Noboru Ohnishi<sup>†‡</sup>

<sup>†</sup>Nagoya University

<sup>†</sup>Aichi Prefectural University

<sup>‡</sup>RIKEN BMC Research Center

上において、照応形「これ」の先行詞は「硫酸第一鉄」でも「蒸留水」でもなく、「硫酸第一鉄を水に溶かしたもの」である。

このような照応を扱うためには、文の表層上の表現だけを指示対象候補とするだけでは不十分であり、文の持つ意味を考慮し、その文によって新たに導入された事象や状態などを推察する必要がある[4]。このような照応を「推論による照応」と呼ぶ。

なお本研究では、名詞類照応で照応形が指示詞または指示連体詞であるものを取り扱うものとする。

## 3 意味を考慮した照応解決

前節で述べたような、推論による照応を解決するためには、動詞の表わす操作による、(指示対象である)名詞の表わす対象の変化を捉えることが必要になる。

このため、中学・高校の理科の教科書にあらわれる文から抽出した約20の例文をもとに分析を行い、それにより状態の変化の推論方法とこの場合に成り立つ制約を考察した。なおこの際、各文は正しく形態素解析をされ、妥当な構文解析結果が得られているものと仮定している。

以下では、分析によって得た、指示対象の状態の変化の推論方法と、そのもとでの照応解決方法について述べる。

## 3.1 処理の手順

照応形の先行詞の同定を、以下のような手続きで行う。

1. 文の意味表現の生成
2. 指示対象候補の変化の推定
3. 先行詞の同定

1.の「文の意味表現の生成」では、各々の文の意味を格フレームによって表現する。本研究の従来との違

いは、2.の「指示対象候補の変化の推定」にある。ここでは、先に生成した意味表現を利用して、指示対象の状態の変化を推定し、その結果を受けて、先行詞の候補のリストの変更を行う。(実際には1.と2.は同時に行われる)最後に、3.で指示詞の先行詞の候補を同定する。

ここで、候補リストの変更において、次のような制約規則を設ける。

#### 前状態不指示の制約：

一般的に、動詞による操作などによって、状態の変化を受けた対象は後文においてその変化前の状態を指示されることはない

この制約は、「談話の一貫性」や「新旧情報」といった立場から導くことができる。また実際に、対象とした例文中にもこの制約を破るものは存在しなかった。

この制約を用いることの利点は、前の状態を無視することによって、先行詞の対象となる候補を削減できること、また、複文においても単文と同様の処理を可能とすることである。

以下では、先にあげた例文を用いて解析の様子を概略する。

#### 3.1.1 文の意味表現の生成

与えられた文が複文の場合には構文解析の結果得られた用言に係るまとまりごとに順次その格フレーム表現を試みる。(連用接続も同様に扱っている)<sup>1</sup>

S1: とる<sub>1</sub>[対象:結晶<sub>1</sub>, 目標:試験管<sub>1</sub>,...]  
 結晶<sub>1</sub>[概念:物質, 名称:結晶, 状態:個体,...]  
 試験管<sub>1</sub>[概念:容器, 名称:試験管,...]

ここまで解析した段階では、指示詞の先行詞の候補リストは、上から

{ 結晶<sub>1</sub>, 試験管<sub>1</sub> }

となる。(先行詞を名詞句に限定していることより)

#### 3.1.2 指示対象候補の変化の推定

ここで、「とる」という用言の辞書に、以下のようなスロット間に関する知識を記述しておく。

とる[状態変化: INC ((対象物質)(目標容器)),  
 MIX ((対象物質)(目標物質))]

この規則を適用することで、先の候補リストの対象は結びつけられ、それらで1つの要素を生成し、また、先の前状態不指示の制約によって、以前の候補はリス

<sup>1</sup>簡単のため、例ではかなり省略した表現を用いた。

トから除かれる。したがって、リスト要素は以下のように変化する。

{ 新対象<sub>1</sub> }

新対象<sub>1</sub>[INC(結晶<sub>1</sub> 試験管<sub>1</sub>)]

これでS1の解析は終了する。

以下順次S2,S3,S4と解析を進めることで、最終的な候補リストは、

{ 新対象<sub>n</sub> }

新対象<sub>n</sub>[INC(MIX(結晶<sub>1</sub> 蒸留水<sub>1</sub>) 試験管<sub>1</sub>)]

のようになる。

#### 3.1.3 指示対象の同定

最後に、最終的な候補リストから照応形の先行詞の同定を行う。候補が複数存在する場合には、意味的な整合性や近接性などを考慮して、もっとも可能性の高いものを候補として同定する。

例文の解析では、候補は1つであるので、これを指示対象とする。この場合、同定された先行詞は正しいものである。

## 4 おわりに

本稿では、限られた場合においてはああるが、文の意味表現を与えることによって、推論による照応の先行詞の同定方法を提案した。提案手法では、文の意味表現とそれによる文中の対象物の状態の変化を考慮することによって、指示対象が直接文中にない場合にも、正しく先行詞の同定ができることを示した。現在、提案した手法に基づいて、与えられた文から意味表現を構築し、先行詞の同定を行うシステムの実装を進めている。今後は、実装したシステムの評価と適用可能な文例範囲の拡大などが課題となる。

## 参考文献

- [1] 山村毅, 大西昇, 杉江昇: “日本語文の照応問題解決システム”, 信学論 (D-II), Vol.J73-D-II No.6, pp.887-896, 1990.
- [2] 山村毅, 大西昇, 杉江昇: “日本語指示詞の前方照応現象の分類”, 信学論 (D-II), Vol.J75-D-II No.2, pp.371-378, 1992.
- [3] 村田真樹, 長尾真: “用例や表層表現を用いた日本語文章中の指示詞・代名詞・ゼロ代名詞の指示対象の推定”, 情処研報, 95-NL-108, pp.43-50, 1995.
- [4] 石崎俊: “自然言語処理 (Natural Language Processing)”, 昭晃堂, 1995.