

指示代名詞の照応問題解決方法についての考察*

4R-2

水野貴文 瀧口伸雄 小谷善行 西村恕彦
(東京農工大学 工学部 電子情報工学科)

1 はじめに

自然言語処理において大きな課題に照応問題解決がある。照応とは文中で同じ名詞が繰り返し登場することを避ける為に代名詞、指示詞などで置き換える言語現象を指す。照応問題解決とはそれらの指示詞の指し示す対象を推論によって解決することである。

本研究では、意味解析レベルを対象し、三段階の推論方法で照応問題解決を行うシステムを設計・実現した。実際の小説中の文章を用いて評価する。

2 深層格モデル上での照応問題解決

ここでは適用する三つの推論方法を例文を挙げながら説明する。

2.1 名詞概念の一致と動詞の直接格支配

照応の対象となる名詞の持つ概念は、指示代名詞にも受け継がれる。例文を示す。

「ぼくは昨日、料理屋へ行った。」 (1)

「そこで、そばを、食べたんだ。」 (2)

文2中の“そこ”から指示対象は場所の属性を持つ名詞とわかる。まずこの手続きにより候補を挙げる。なお、本研究で使用した名詞概念[2]は“具体名詞”、“現象名詞”、“抽象名詞”、“制限緩やか”、をそれぞれ上位概念とした計19種の属性である。次に動作フレーム(動詞)の格支配を直接受けているものに評価値を加える。評価値については2.4節で詳しく述べる。例を挙げる。

「M市にあるBホテルに行け。」 (3)

「そのロビーで指示を待て。」 (4)

例文はある短編推理小説¹の一会話である。候補として“M市”と“Bホテル”が挙がる。動詞(動作フレーム)“行け”に直接格支配されているものの評価値を上げる。文3が「BホテルのあるM市に行け」という形

*Resolving Anaphora Reference Problems of Demonstrative Pronouns, Takafumi MIZUNO, Nobuo TAKIGUCHI, Yoshiyuki KOTANI, Hirohiko NISIMURA Dept. of Computer Science, Tokyo University of Agric. and Tech.

¹山村正夫編:“新トリックゲーム”:日本文芸社(1988)から引用、一部改変した。

であれば次にくる文章の“そこ”は“M市”を示すのが自然である。

2.2 動詞間の意味接続

2.1節で挙げられた候補に対して、動詞間の意味接続や因果規則を適用して候補を絞る。因果規則を適用し、意味的に解決する方法は文献[1]で有用性が報告されている。例を挙げる。

「うっかり、花瓶を机から落としたんだ。」(5)

「それは、粉々に、こわれたよ。」(6)

文5と文6には動詞間に意味的につながりがあると考えられる。“それ”は“こわれた”対象であり、指示対象は前文の“落とした”対象である。この知識をシステムが知っていれば推論できるが、すべての動詞の組み合わせを考えると膨大な量になる。そこで動詞の意味分類[2]を使い、組み合わせ数を減らす。意味分類は、状態、存在・所有、接触、生理・心理といった28種の要素の組み合わせで272通りの意味を表わした。文5と文6について手順を下に述べる。

1. “それ”を含む動作フレーム“こわれる”を得る
2. “こわれる”の意味分類Aを得る
3. その動作フレームの動作前スロット“落とす”を得る
4. 動作前スロットがなければ、終了する8.へ
5. “前動作フレーム”(一回目なら“落とす”)の意味分類Bを得る
6. 意味分類A、Bより意味接続規則辞書を参照し指示対象のスロット名が存在すれば8.へ
7. 規則になれば、前文動作フレーム“落とす”の更に前動作フレームを得て5.へ
8. 推論を終了する

意味接続規則辞書には“状態変化、消滅した(こわれた)対象は、状態変化、移動した(落とした)対象である”という知識が記述されている。

2.3 意味解析レベルにおける焦点抽出

会話で場所が移り変わる場合、最後にでてきた場所名詞が指示対象になることが多い。知識フレームで言えば、指示代名詞のフレームに対して距離的に一番近いものを候補とする。例を挙げる。

「家から料理屋へ行った。

そこで学生達を見た。」 (7)

動作フレーム“行った”の起点スロットは“家”、終点スロットは“料理屋”である。手順を下に示す。

1. 指示詞の動作フレームを参照し“動作前、状態前スロット”から前動作フレームを得る
2. その動作フレームに終点スロットがあればそのフレームを焦点とし7.へ
3. その動作フレームに場所スロットがあればそのフレームを焦点とし7.へ
4. その動作フレームに動作前スロットがあればそのフレームを得て2.へ
5. その動作フレームに状態前スロットがあればそのフレームを得て2.へ
6. 2.～5.に失敗したら終了する
7. 推論を終了する

2.4 システムの適用順序

2.1節～2.3節に示した推論法の利用に付いて述べる。動詞間の意味接続は場所スロットに関して全く無知である。図1に入力から出力までの流れを示す。括弧内の数値は各推論法が適用できたときに加える評価値である。すべての推論規則の適用を終了した時点で最も評価が高いものを指示対象とし、フレームをつなぎ変える。

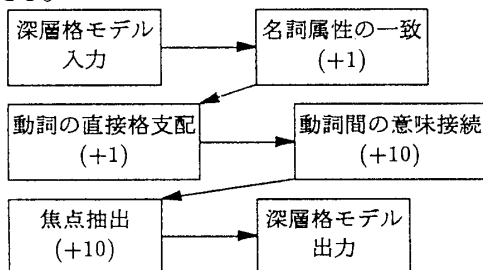


図1: 推論システムの流れと加える評価値

3 実験と考察

2章で述べた推論を行うシステムを作成し、実験を行った。小説から指示代名詞が現われる会話文13組に

対し12組が正確に推論できた。文5と文6を例に挙げる図2をシステムに入力すると中間形式は図3のようになり評価値が最も多い“花瓶”のフレームと“それ”がリンクし、終了する。指示代名詞の照応問題解決について述べたが、以上の推論規則は比較的拡張も、容易で有効であるといえる。

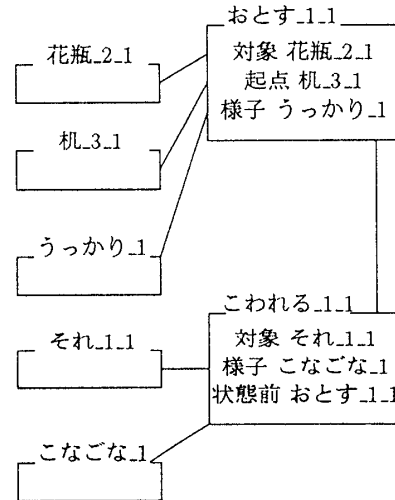


図2: 入力モデル

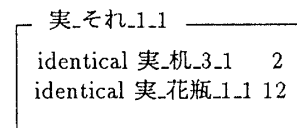


図3: 中間形式の指示詞フレーム

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, 電子情報通信学会論文誌, 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, 電子情報通信学会論文誌, J75-D-II No.2, pp.371-378(1992)
- [3] 情報処理振興事業協会技術センター: 計算機用日本語基本動詞辞書 IPAL(BasicVerbs) - 解説編 -, 情報処理振興事業協会 (1987)