

# 副詞的使用に着目した「名詞+に」の深層格識別法

永田 拓<sup>†</sup> 石川 勉<sup>†</sup>  
 拓殖大学工学部情報工学科<sup>†</sup>

## 1 はじめに

我々は、「言葉で考えるコンピュータ」の実現を目指して研究している [1]. このコンピュータは、自然言語文で入力された質問に対し、常識知識と概略推論 [2] をベースとした推論法を用いて解を導き出す. ここの知識表現は、言葉をベースとした述語論理を拡張した形式であり、引数には基本的に深層格をとる [3]. 推論処理では、この深層格ごとに照合(単一化)を行うため、自然言語文からこの形式に変換する際、格助詞に基づく深層格の判別が重要となる. 特に格助詞の中でも“に”は、場所格、時間格、目標格、副詞格など多くの格で用いられる.

本稿では、この中から副詞格の判別に着目し、その判別手法について提案する. 具体的には、「名詞+“に”」で副詞となる語(以下、副詞派生語)を対象とし、解析結果から得られる名詞の品詞情報やシソーラス情報等を用いた判別法を示す.

## 2 知識表現形式と知識変換の流れ

### 2.1 想定している知識表現法

単文、複文に係わらず一つの文は一つの述語式(素式)で表す. その形式を以下に示す.

$$sP(r_1:t_1, r_2:t_2, \dots, r_n:t_n)$$

ここで、s は様相子(過去、否定等)、P は述語記号であり、 $t_i$  は引数、 $r_i$  はそのラベル(EDR 電子化辞書 [4] の深層格を基に設定)である. これら述語記号、引数は基本的に全て言葉を用いる. 例えば、「太郎は絶対に五時に自宅に帰る。」という文は以下のような表現となる.

例) 太郎は絶対に五時に自宅に帰る.  
 帰る (agt:太郎,gol:自宅,tme:五時,mnr:絶対)

ここで、agt,gol,tme,mnr はそれぞれ主格、目標格、時間格、副詞格を表すラベルである. 複文の場合には主節が述語式の基本を構成し、従属節は述語部あるいは引数部に埋め込まれて表現される [3]. 推論処理では、この表現は順序ソート論理に変換され、同じラベルの引数間で単一化が行われる [5]. 本稿では、記述の簡略化のため変換前の形式で説明する.

### 2.2 知識変換の流れ

知識変換の流れを前節の文を例に図 1 に示す. まず、文を形態素・構文解析(それぞれ茶釜 [6], 南瓜 [7] を利用)し、単語の品詞情報(品詞体系、活用型、活用形)、文節の係り受けおよび格情報を獲得する. それらを基に必須格、任意格をラベル付けし、述語式に変換する. ここで必須格のラベル付けには、EDR 電子化辞書の日本語動詞共起パターン副辞書(以下、共起辞書)および日本語単語辞書(以下、単語辞書)を利用する. 共起辞書では、動詞毎に、必須格のラベル、それに対応する格助詞およびその格が取り得る概念の意味的なカテゴリを表す概念識別子が登録されている. まず、これらと解析

結果の情報を照合し、深層格の候補を決定する. 次に、格助詞の前方の名詞が先の候補の意味カテゴリの範囲に含まれるか否かを単語辞書を用いてチェックし、最終的に深層格を決定する. その後、必須格として抽出されなかった格を任意格とし、解析結果の品詞情報やカテゴリ情報等を基にラベル付けを行う.

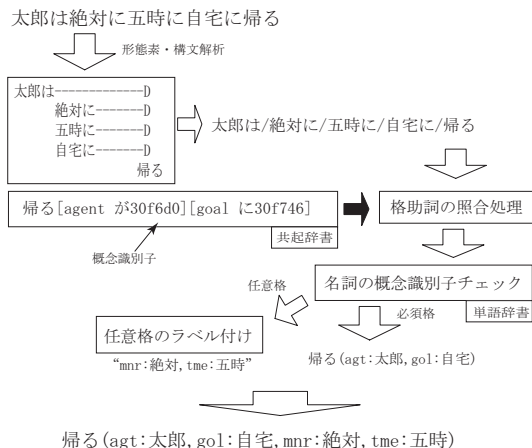


図 1: 知識変換の流れ

## 3 副詞判別法

### 3.1 任意格のラベル付け

深層格を的確に判別することは極めて難しい. 例えば EDR 電子化辞書のコーパスでは、「木々が湖水に映る」の“湖水に”を gol と、また「男が敵を恨む」の“敵”も gol としているが、これらはそれぞれは plc(場所格),obj(目的格)とも言える. すなわち、深層格自身の意味的な定義が必ずしも明確にできないため、人手でもその判別が曖昧となる. 従って、我々は各格助詞毎に深層格の意味が明確で、かつ技術的に精度良く判別できるものから順に判別していくアプローチをとっている. 具体的には、格助詞“に”には次のような格があるが、現在この格については上から順に優先的に決定している. 副詞格の判定を最初に行っているのは、この格は個体を示すのではなく、述語部を単に修飾する格であることも一つの理由である.

- 副詞格 (mnr) 例) 太郎は堅実に金を貯める.
- 場所格 (plc) 例) 太郎は東京に家を建てる.
- 時間格 (tme) 例) 太郎は仕事を年末に残す.
- 目標格 (gol) 例) 太郎は荷物を東京に送る.
- 場面格 (scn) 例) 太郎は会議に参加した.
- 原因格 (cas) 例) 太郎は恐怖に震えた.
- 目的格 (pps) 例) 太郎は本を返しに行く.

### 3.2 副詞判別手順

副詞判別の流れを図 2 に示す. 副詞格は、基本的に解析結果から得られた格助詞“に”の前方の名詞の品詞体系と 3.3 節で述べる副詞判別条件を基に判別を行う. 判別する名詞によっては、時間カテゴリも利用する.(理由は 3.3 節で述べる.) 副詞判別条件で判別できない語については、3.4 節で述べる副詞の使用が可能と判断できる名詞を集めたテーブルを利用して、副詞判別を行

A Judging Method for Phrase "Noun + ni" Used as Adverb

<sup>†</sup> Taku Nagata, Tsutomu Ishikawa

<sup>†</sup> Department of Computer Science, Takushoku University

う。これらの処理で、副詞と判断されなかった語については、上述した方法で任意格のラベル付けを行う。2.1節の文を例に具体的に説明すると、名詞“絶対”の品詞体系「名詞 副詞可能」は判別条件と一致し、時間のカテゴリに“絶対”は含まれないので、mnrのラベルを付与する。一方、名詞“時”の品詞体系「名詞 接尾 助数詞」は判別条件と一致せず、副詞テーブルに登録してある名詞と“時”が一致しないため、他の任意格のラベル付けを行う。

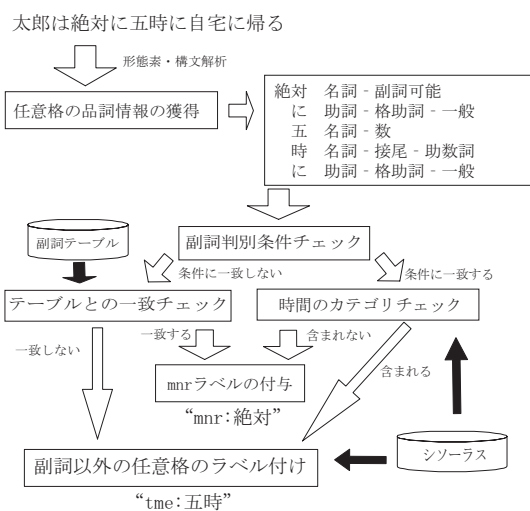


図 2: 副詞判別の流れ

### 3.3 副詞判別条件

副詞判別法は、格助詞“に”の前方の名詞（以下、副詞派生名詞）の品詞体系に着目し、「副詞派生名詞の品詞体系+“に”」の形式を副詞判別の条件とする。現代副詞用語辞典 [8]（以下、副詞辞典）を基に調査し、以下の3つの形式を副詞として判別する条件に設定した。

- ① 「副詞+“に”」 例) 実際に
- ② 「名詞 形容動詞語幹+“に”」 例) 完璧に
- ③ 「名詞 副詞可能+“に”」 例) 一遍に

しかし、条件③の場合は時間格もありうる。例えば、「太郎は夕方に自宅に帰る」という文を変換すると“夕方に”の品詞体系が条件③と一致してしまい、本来、tmeでラベル付けされるべきところを、mnrでラベル付けする誤変換が生じる。そこで、シソーラス [9] を利用し、副詞派生名詞が時間のカテゴリに含まれているかをチェックし、含まれていない場合のみ、副詞と判別する。

### 3.4 副詞テーブル

副詞派生語の中には、上述した副詞判別条件の形式以外にも「名詞 一般+“に”」（一身に）や「名詞 サ変接続+“に”」（一緒に）等が存在する。しかし、これらの形式は副詞だけを表すとは限らないため、判別条件に設定することができない。例えば、“旅に”という語を解析すると“旅”の品詞体系が「名詞 サ変接続」となり、副詞派生語“一緒に”と同じ形式になってしまう。そこで、「品詞体系+“に”」の形式では判別できない副詞派生名詞を集め、副詞テーブル（以下、テーブル）を作成し、これを利用して副詞判別を行う。

具体的には、テーブルに副詞派生名詞の品詞体系ではなく、副詞派生名詞そのものを登録しておき、格助詞が“に”であり、その前方の名詞がテーブル中の副詞派生名詞と一致した場合、副詞として強制的にラベル付けを行う。

さらに、「\*”的+“に”」（積極的に）や「\*”気味+“に”」（興奮気味に）など定型的につかう形式にもテーブルで対応する。これらの場合、“\*”の部分があるような品詞体系の単語であろうとも副詞であると判断できるため、“的”、“気味”等をテーブルに登録することで副詞判別を可能としている。現在、テーブルには副詞辞典を基に“一目散”、“存分”など45個の副詞派生名詞を登録している。

## 4 評価

上述の副詞判別法を用いて、新聞データから抽出した副詞が含まれている200文、副詞が含まれていない100文、計300個の単文を対象に知識変換を行った。

その結果、副詞が含まれている200文すべてにおいて副詞を判別することに成功した。また、副詞を含まない100文から文中の名詞を副詞として誤変換する文は一つも発見されなかった。各条件によって判別された副詞派生語を以下に示す。

- 条件① : すぐに、現に、にわかに、実際に、次々に
  - 条件② : 意外に、正式に、急速に、余分に、頻繁に
  - 条件③ : 絶対に、一斉に、真っすぐに
- テーブル: 得意そうに、真っ先に、爆発的に、口々に

## 5 まとめ

助詞“に”の直前の名詞の品詞情報、シソーラスの時間カテゴリ、および副詞テーブルを利用することにより、「名詞+“に”」の副詞的使用を正確に判別する手法を提案した。この手法を、新聞データから抽出した300文に適用した結果、この種の副詞を的確に判別することに成功した。

## 参考文献

- [1] 石川勉: 言葉で考えるコンピュータの実現にむけて、人工知能学会論文誌 (第17回) 論文集, 3C1-04(2004).
- [2] Nguyen Viet Ha, 石川勉, 阿部明典: 知識の類似性を利用した概略推論法, 電子情報通信学会論文誌 DI, Vol. J84-D-1, No. 4, pp. 389-400(2001).
- [3] 北野正樹, 石川勉: 言葉をベースとした述語形式の知識表現法と順序ソート論理によるその推論処理法, 人工知能学会, ことば工学研究会資料 SIG-LSE-A503, pp. 43-52(2006).
- [4] 日本電子化辞書研究所: EDR 電子化辞書, 日本語動詞共起パターン副辞書, 日本語単語辞書 (1996).
- [5] M. Kitano, S. Nishita, T. Ishikawa: Inference Scheme for Order-Sorted Logic Using Noun Phrases with Variables as Sorts, 5th Mexican International Conference on Artificial Intelligence (MICAI2006), Springer-Verlag LNAI 4293, pp. 49-58(2006).
- [6] 茶釜: <http://chasen.naist.jp/hiki/ChaSen/>
- [7] 南瓜: <http://chasen.org/taku/software/cabocho/>
- [8] 飛田良文, 浅田秀子: “現代副詞用語辞典, 東京堂出版”
- [9] 池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林: “日本語語彙大系, 岩波書店”