

話者の対象認識過程を分析するパーザの基本的枠組み

3R-3

沼崎 浩明 高橋 大和 宮崎 正弘

新潟大学・大学院工学研究科

1 はじめに

本稿では、DCG 文法を一般化 LR 法に基づいて構文解析する SGLR パーザ [1] の上に、日本語の話者の対象認識過程を分析する機構を組み込むための、基本的枠組について提案する。そして、その動作を計算機上で確認した結果について述べる。話者の対象認識過程とは、対象を認識し、それを言語として表現する対象を概念化し、対象に対する話者の見方や捉え方、判断等を識別する過程のことをいう。筆者らは、特に、三浦文法に基づいて考案された日本語の助詞「が」と「は」についての意味規則 [2] を用い、その規則の動作機構を DCG の補強項で実現して SGLR パーザで実行できるようにしている。

2 助詞「が」と「は」の意味規則

三浦文法に基づく助詞「が」と「は」の意味規則は、既に文献 [2] によって提案されている。その要点は次の通りである。まず、意味を「音声や文字の種類に結び付き固定された対象と認識との間の関係」即ち、対象と認識との間の関係と定義した上で、助詞「が」と「は」に前接する名詞の3つの範疇（クラスとインスタンスを表す範疇N1、インスタンスを表す範疇N2、クラスを表す範疇N3）に対して助詞「が」と「は」の意味分類を次のように定義している。

[N1+が] 文脈よりN1がN3と捨象される場合、限定性。前接する名詞が目的格の場合、N1をN3と捨象し限定性。
例：酒が好きだ。／水が飲みたい。

前接する名詞が総記の場合、N1をN3と捨象し限定性。

例：燕が鳥だ。／子供がかかりやすい。

上記以外の場合、N1をN2と捨象し個別性。

例：鳥が飛ぶ。／雪が白い。／犬がいる。

[N1+は] 文脈よりN1がN2捨象される場合、特殊性。

存在文の場合、N1をN3と捨象し特殊性。

例：犬はいる。／本はある。

前接する名詞が対照の場合、N1をN3と捨象し特殊性。

例：月は東に日は西に。

前接する名詞が目的格の場合、N1をN3と捨象し特殊性。

例：酒は好きだ。／水は飲みたい。

上記以外の場合、N1をN3と捨象し普遍性。

例：鳥は飛ぶ。／雪は白い。／燕は鳥だ。

[N2+が] 限定性。例：太郎が学生です。

[N2+は] 特殊性。例：太郎は学生です。

[N3+が] 限定性。例：鳥類がハチュウ類から進化した。

[N3+は] 普遍性。例：鳥類はハチュウ類から進化した。

3 パーザの基本的枠組

次に話者の対象認識過程を分析するパーザの基本的枠組について記述する。

・文法規則は DCG 形式とする。

・DCG の記述はチョムスキー標準形とする。

・全ての名詞に（引数として）N1、N2、N3、N4、N5 の分類を与える（ただし、N4 は動作名詞、N5 は状態名詞を表す）。

・全ての助詞、助動詞に（引数として）その語を与える。

・全ての動詞に、上位の意味概念を与える [2] の意味分類では用いられないものも含む。

・意味分類は補強項で行なう。

・構文解析と意味分類を同時に進める。

・意味分類の動作は、トップレベルの規則で呼び出す。

・構文的曖昧性がある場合は、各解析に意味分類を与える。

4 文法と辞書の試作

3に基づき試作した簡易版の文法と辞書を示す。

文(S)→後置詞句(PP), 文(S1), {分類([PP, S1], S)}.

文(S)→後置詞句(PP), 動詞(V), {分類([PP, V], S)}.

文(S)→後置詞句(PP), 形容詞(ADJ), {分類([PP, ADJ], S)}.

文(S)→後置詞句(PP), 述語(PRED), {分類([PP, PRED], S)}.

文(S)→文(S1), 文(S2), {分類([S1, S2], S)}.

後置詞句([N, P])→名詞(N), 助詞(P).

後置詞句([PP1, PP2])→後置詞句(PP1), 後置詞句(PP2).

述語([N, V])→名詞(N), 形式動詞(V), {N=n4}.

述語([V, P])→動詞(V), 助動詞(P).

述語([V, ADJ])→動詞(V), 形式形容詞(ADJ).

述語([N, da])→名詞(N), 助動詞(da).

述語(□)→□.

名詞(n1)→[鯨]. 名詞(n1)→[犬]. 名詞(n1)→[学生].

名詞(n1)→[雨]. 名詞(n1)→[本]. 名詞(n1)→[子供].

名詞(n1)→[酒]. 名詞(n1)→[月]. 名詞(n1)→[雪].

Basic Form of Parser for Analyzing Viewpoint of Speaker's Recognition of Japanese Particle "ga" and "ha".
Hiroaki NUMAZAKI, Yamato TAKAHASHI, Masahiro MIYAZAKI
Nigata University

名詞 (n1)-->[水]. 名詞 (n1)-->[東]. 名詞 (n1)-->[西].
 名詞 (n1)-->[燕]. 名詞 (n1)-->[日]. 名詞 (n1)-->[鳥].
 名詞 (n2)-->[私]. 名詞 (n2)-->[僕]. 名詞 (n2)-->[太郎].
 名詞 (n2)-->[花子]. 名詞 (n3)-->[鳥類].
 名詞 (n3)-->[ハチュウ類]. 名詞 (n3)-->[ホニユウ類].
 名詞 (n4)-->[進化]. 名詞 (n5)-->[好き].
 助詞 (ha)-->[は]. 助詞 (ga)-->[が]. 助詞 (ni)-->[に].
 助詞 (kara)-->[から].
 助動詞 (da)-->[です]. 助動詞 (da)-->[だ].
 助動詞 (tai)-->[たい]. 助動詞 (masu)-->[ます].
 動詞 (現象)-->[かかり]. 動詞 (現象)-->[降って].
 動詞 (行為)-->[飛ぶ]. 動詞 (行為)-->[飲み].
 動詞 (行為)-->[行く]. 動詞 (存在)-->[いる].
 動詞 (存在)-->[ある].
 形式動詞 (存在)-->[い]. 形式動詞 (行為)-->[した].
 形容詞 (色)-->[白い].
 形式形容詞 (程度)-->[やすい].

次に、意味分類の処理規則を示す。

分類 ([[n1,ga],[n1,da]], [[N 1→N 3+が]], 総記, 限定性):-!.
 分類 ([[n1,ga],[_,da]], [[N 1→N 3+が]], 目的格, 限定性).
 分類 ([[n1,ga],[V,tai]], [[N 1→N 3+が]], 目的格, 限定性]).
 分類 ([[n1,ga],[V,程度]], [[N 1→N 3+が]], 総記, 限定性):-!.
 分類 ([[n1,ga],[_],[N 1→N 2+が]], 中立叙述, 個別性).
 分類 ([[n1,ha],存在],[[N 1→N 3+は]], 存在, 特殊性).
 分類 ([[n1,ha],[_,ni]], [], [[N 1→N 3+は]], 存在, 特殊性]).
 分類 ([[A,B,C],[A,B,C]], [A,対照,C]).
 分類 ([[n1,ha],[n4,da]], [[N 1→N 3+は]], 目的格, 特殊性]).
 分類 ([[n1,ha],[n5,da]], [[N 1→N 3+は]], 目的格, 特殊性]).
 分類 ([[n1,ha],[_,tai]], [[N 1→N 3+は]], 目的格, 特殊性]).
 分類 ([[n1,ha],[_],[N 1→N 3+は]], 総称, 普遍性).
 分類 ([[n2,ga],[_],[N 2+が]], 総記, 限定性).
 分類 ([[n2,ha],[_],[N 2+は]], 文脈指示, 特殊性).
 分類 ([[n3,ga],[_],[N 3+が]], 総記, 限定性).
 分類 ([[n3,ha],[_],[N 3+は]], 総称, 普遍性).

5 実験結果

上記の文法、辞書、意味分類規則を SGLR パーザ [1] 上で動作させた結果を下に示す。

```
input sentence : 酒, が, 好き, だ.
Length : 4
execution time      = 0 msec
|-文
  |-後置詞句
  |   |-名詞 -- 酒
  |   |-助詞 -- が
  |-述語
  |   |-名詞 -- 好き
  |   |-助動詞 -- だ
```

Argument Information: [[[N 1→N 3+が]], 目的格, 限定性]]
 Number of Trees is : 1

input sentence : 月, は, 東, に, 日, は, 西, に.

```
Length : 8
execution time      = 20 msec
|-文
  |-文
  |   |-後置詞句
  |   |   |-後置詞句
  |   |   |   |-名詞 -- 月
  |   |   |   |-助詞 -- は
  |   |   |-後置詞句
  |   |       |-名詞 -- 東
  |   |       |-助詞 -- に
  |   |-述語 -- []
  |-文
  |   |-後置詞句
  |   |   |-後置詞句
  |   |   |   |-名詞 -- 日
  |   |   |   |-助詞 -- は
  |   |   |-後置詞句
  |   |       |-名詞 -- 西
  |   |       |-助詞 -- に
  |-述語 -- []
```

Argument Information: [[[N 1→N 3+は]], 対照, 特殊性]]
 Number of Trees is : 1

上記のように本稿に記載した例文については、全て正しく意味分類がなされた。また、予想されていたことではあるが、構文的曖昧性も著しく減少することが判明した。例えば、「月は東に日は西に。」の文には、55 通りの構文木が存在するが、意味制約によりそれが 1 つに絞られ、解析時間も 3 倍速くなっている。

6 結論

本研究では、DCG 文法に基づいて、助詞「が」と「は」の意味分類を行なうパーザの基本的枠組を提案し、その有用性を実証した。今後は、三浦文法による形態素解析システム [3] と結合し、三浦文法による本格的な統語意味融合型の文法を構築し、構文的曖昧性の意味制約を用いた解消の検討を行なう予定である。

参考文献

- [1] 沼崎浩明, 田中穂積: SGLR: 逐次型一般化 LR パーザの Prolog による実現, 情報処理学会論文誌, Vol.32, No.3, pp.396-403 (1991)
- [2] 宮崎正弘, 高橋大和: 話者の対象認識過程からみた助詞「が」「は」の意味分析, 情報処理学会第 46 回全国大会, No.1B-8 (1993)
- [3] 高橋, 佐野, 宍倉, 前川, 宮崎: 頑健性を目指した日本語形態素解析システムの試作, 自然言語処理シンポジウム (自然言語処理における実動) 論文集 (1993)