

自律オブジェクトによる右隅統語解析

笠 晃一†

本田 幸司†

横田 将生††

† ㈱日本データベースネットワーク研究所

†† 福岡工業大学

文脈自由文法を用いて自然言語文の統語解析を行なおうとする場合、無駄な解析木が多量に生成されることが知られている。このような無駄な解析木の生成は、到達可能性や接続可能性を用いて、ある程度まで抑制することができる。このうち、到達可能性は種々の統語解析アルゴリズムで採用され、それなりの効果を挙げているようである。しかしながら、到達可能性や接続可能性を用いても、依然としてかなりの量の無駄な解析木が残り、解析速度や消費記憶域に悪影響を与えている。我々は、自律オブジェクトを用いた統語解析法を開発したが、これは大域的制約を用いているため、無駄な解析木の生成をほぼ完全に抑制することができる。また、この方法は、オブジェクト指向の考えを取り入れているため、自然言語のもつ多様さや複雑さを系統的に記述することもできる。

Right-Corner Parsing Method with Autonomous Objects

Koichi Ryu†

Koji Honda†

Masao Yokota††

† Japan Database Network Laboratory Co., Ltd.

†† Fukuoka Institute of Technology

It is well known that many useless parsing trees are generated when natural language sentences are parsed with a context-free grammar. Reachability or connectability can suppress the generation of useless trees to some extent. The reachability has been used in various parsing algorithms and seems to be rather effective. However, even if reachability or connectability is used, there are still a considerable number of useless trees left and they unfavorably affect parsing speed or consumed memories. We have developed a parsing method with autonomous objects, which can suppress the generation of useless trees almost completely. This method also makes it possible to describe the variety or complicatedness of natural languages systematically, because the idea of object oriented programming is incorporated in it.

1. はじめに

自然言語文は、文脈自由文法を用いて解析されることが多いが、たいいていの場合、無駄な部分解析木（あるいは、部分解析木に対応する状態）が多量に発生することが知られている。そして、特に次のような考えに基づいて作られた統語解析システムの場合、このような無駄な部分解析木はできるだけ少ない方が望ましい。

(1) 単一化に基づくシステム

最近、単一化に基づいて統語解析を行なうシステム^{[1][2]}が数多く発表されるようになってきた。単一化を用いると、自然言語に関する知識を簡潔に表現できたり、統語解析と意味解析を融合して組合せの爆発を回避できたりするからである。しかし、単一化という操作は、かなりの時間を消費するので、これらのシステムの中には、いかにして単一化の回数を削減するかという工夫を行なっているものが少なくない。この場合、無駄な部分解析木を削減すれば、単一化の回数も削減できることは明らかである。

(2) オブジェクト指向に基づくシステム

オブジェクト指向に基づいて統語解析を行なうシステム^{[3][4]}も多くなってきた。この考え方をを用いると、自然言語のもつ階層性や特殊性をうまく記述できるからである。しかし、オブジェクトを生成するのに必要な時間は無視できないものである。特に、一つの部分解析木を一つのオブジェクトに対応させる方法の場合、部分解析木の生成をできるだけ少なくする必要がある。ここでも、無駄な部分解析木を削減することが有効になる。

そこで、このような無駄な部分解析木の発生を防ぐための方法がいくつか考えられており、たとえば、代表的なものとして、到達可能性の利用と接続可能性の利用を挙げることができる。このうち、到達可能性は、種々の統語解析アルゴリズム^{[5][6][7]}で採用されており、欧文の解析に対しては確かに効果があるように思われる。接続可能性も、数は少ないが、いくつかのシステム^[8]で利用されている。ただし、いずれの方法も局所的な制約しか利用していないので、依然として、かなりの量の無駄な部分解析木が残ってしまうという欠点があった。

これに対し、我々は、大域的な制約を利用

するような統語解析法を開発し、無駄な部分解析木をほとんど生成しないようにすることができた。これは、単一化とオブジェクト指向の両方に基づくシステムであるが、両方の特長を兼ね備えているだけでなく、高速な解析も可能になっている。

2. 到達可能性と接続可能性

到達可能性を利用すれば、トップダウン的な予測により、適用可能な文法規則を絞り込むことができ、これによって無駄な部分解析木を削減することができる。また、接続可能性を利用すれば、文法範疇と文法範疇がつながるかどうかを調べることにより、同様なことができる。

2. 1. チャート統語解析法

本論文の記述は、すべて、チャート統語解析法^[9]の用語を用いて行なうことにする。チャート統語解析法は、統語解析のための枠組みを示すものなので、今の目的には好都合だからである。

チャート法では、部分的な統語解析木を表すのに項と呼ばれるデータ構造を用いる。たとえば、下の(項1)は項であり、これは、図1のような部分解析木を表している。

(項1) $[[[\text{日本人}]_n]_{n_p} [?]_p]_{p_p}$

ここに、 $[?]_p$ は、まだ内容が決まらない後置詞(p)を表わす。

項には、 $[[[\text{日本人}]_n]_{n_p}]$ のように?を含まないものと、 $[[[\text{日本人}]_n]_{n_p} [?]_p]_{p_p}$ のように?を含むものがあり、それぞれ、完成した部分解析木と未完成の部分解析木とに対応している。チャート法では、解析の状態をグラフで表現するので、これらは、それぞれ、不活性弧、活性弧と呼ばれることもある。グラフ上では、不活性弧を実線で、活性弧を破

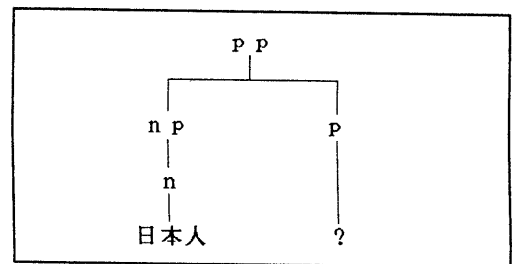


図1. 項1が表す部分解析木

線で表すことにする。

2. 2. 到達可能性

到達可能性^[10]は、次のように再帰的に定義することができる。ただし、 $C_1 \sim C_3$ は文法範疇である。

[定義 1] 到達可能性

- ① C_1 を右辺の先頭にもち、 C_2 を左辺に持つような文法規則があれば、 C_1 から C_2 へ到達可能である。
- ② C_1 から C_2 へ到達可能であり、 C_2 から C_3 へ到達可能であれば、 C_1 から C_3 へ到達可能である。

たとえば、図 2 のような文法規則を考えてみよう。この文法規則に対して、表 1 のような到達可能性表を作成することができる。これは、表の左側にある文法範疇が、表の右側にある文法範疇に到達可能であることを表している。すべての文法範疇は、自分自身にも到達可能であるが、このような関係は表には書き入っていない。

次に、この到達可能性表を使って、無駄な部分解析木がどのようにして削除されるかを

ap → a	n → 日本人
ap → adv ap	n → 神経
np → n	adv → 実に
np → ap np	a → 細やかな
np → s np	p → は
pp → np p	p → を
vp → v	v → 持つ
vp → vp auxv	auxv → ている
s → vp	
s → pp s	

図 2. 文法規則の例 (その 1)

a	ap np pp s
adv	ap np pp s
ap	np pp s
np	pp s
pp	np s
n	np pp s
v	vp np pp s

表 1. 到達可能性表

見てみよう。例文として、次の (文 1) のようなものを考える。

(文 1)

日本人は実に細やかな神経を持っている

到達可能性を利用しないときは、下の (項 2) のような、無駄な部分解析木が作られる。

(項 2)

[[[細やかな]_a][_{np}][神経]_n]_{np}]_{np}

さて、図 2 の文法規則には、次の (規則 1) のようなものが含まれている。

(規則 1) ap → adv ap

そして、この規則の右辺は、「実に」という副詞 (adv) の次には、形容詞句 (ap) が来ることを予測しているのである。ところが、(項 2) に対応する文法範疇は、名詞句 (np) であって、しかも、到達可能性表を用いると、np から ap へは到達可能でないことが分かる。言い換えると、np が成長して、将来 ap になることは決してないのである。そこで、(項 2) に対応する下の (規則 2) を適用する前に、到達可能性を調べれば、無駄な部分解析木である (項 2) は生成されないことになる。

(規則 2) np → ap np

2. 3. 接続可能性

接続可能性^[10]も、再帰的に定義することができるが、その前に、右到達可能性という概念を導入する。これに対し、従来の到達可能性を、左到達可能性と呼ぶことにしよう。以下の定義において、 $C_1 \sim C_4$ は文法範疇を表すものとする。

[定義 2] 右到達可能性

- ① C_1 を右辺の右端にもち、 C_2 を左辺に持つような文法規則があれば、 C_1 から C_2 へ右到達可能である。
- ② C_1 から C_2 へ右到達可能であり、 C_2 から C_3 へ右到達可能であるならば、 C_1 から C_3 へ右到達可能である。

この定義を用いると、接続可能性を以下のよ

うに定義することができる。

[定義3] 接続可能性

- ① C₁ と C₂ をこの順に連続して右辺に含むような文法規則があるとき、C₂ は C₁ に接続可能である。
- ② C₂ が C₁ に接続可能であるとする。このとき、C₃ から C₁ に右到達可能であり、C₄ から C₂ に左到達可能であるならば、C₄ は C₃ に接続可能である。

ただし、単に接続可能性といったときは、ある文法範疇にどのような語彙範疇が接続するかということのみを問題にすることが多く、ここでも、これより先、この意味で使うことにする。

たとえば、図2のような文法規則に対して、接続可能性表を作成すると、表2のようになる。これは、表の左側にある文法範疇に、表の右側にある語彙範疇が接続可能であることを表している。

ap	a adv n v
np	p
pp	a adv n v
vp	a adv n v auxv
s	a adv n v

表2. 接続可能性表

ここで、例文として再び、(文1)を取りあげ、無駄な部分解析木を削除するのに接続可能性表がどのように使われるかを見てみよう。接続可能性表を利用しないときには、次の(項3)のような無駄な部分解析木が作られる。

(項3) [[[持つ]_v]_{vp}]_s

ところが、一つ先読みをすれば、「持つ」という動詞(v)の後には、「ている」という助動詞(auxv)があることが分かる。そして、接続可能性表を使えば、sにauxvは接続しないことが分かるのである。したがって、(項3)に対応する下の(規則3)を適用する前に、接続可能性を調べれば、無駄な部分解析木である(項3)が生成されることはない。

(規則3) s → vp

2.4. 局所的制約と大域的制約

到達可能性の利用にしろ接続可能性の利用にしろ、結局は局所的な制約を用いている。つまり、到達可能性の利用は、ある文法範疇のすぐ右に来るべき文法範疇を予測し、これによって適用可能な文法規則を絞り込もうとするものである。また、接続可能性の利用は、一つ先読みして次の語彙範疇を知ることにより、このすぐ左に来るべき文法範疇を予測し、適用可能な文法規則を絞り込もうとするものである。しかしながら、このような局所的な制約だけを使っているだけでは削除できないような部分解析木も存在する。

たとえば、(文1)の解析において、次のような無駄な部分解析木が作られる。

(項4)
 [[[[[神経]_n]_{np}]_p]_{pp}][[[[持つ]_v]_{vp}
 [ている]_{auxv}]_{vp}]_s]

これは、名詞(n)である「神経」に「実に細やかな」という修飾句が付くことなしに、「神経を持っている」だけで文(s)が構成されてしまったのである。なぜこのようなことが起こるのであろうか。それは、「神経」と「実に細やかな」という修飾句の関係を見ているときには、「実に細やかな」という修飾句の修飾先が「神経」だけであるとは限定できないからである。つまり、「神経」という名詞の後方にも別の名詞が出現し、「実に細やかな」という修飾句はそちらを修飾することも知れないのである。このようなわけで、「実に細やかな」と「神経」を切り離すような解釈もとっておく必要が生じてしまう。

このことは、逆に考えれば、「神経」という名詞の後に別の名詞が出現しないのであれば、「実に細やかな」という修飾句とその後続く「神経」を切り離す必要はないということになる。そして、今の場合、「神経」という名詞の後には、別の名詞は出現していないので、まさにこの条件を満足するのである。ところで、ある名詞の後方に別の名詞が出現するかどうかというのは、大域的な制約である。つまり、大域的な制約を使って初めて削除可能な部分解析木も存在することが分かる。以下では、まず、大域的な制約を使いやすくするために右隅統語解析法というものを導入し、さらに、無駄な解析木を削除するための統一的な方法として消去可能性という概念を説明する。

3. 右隅統語解析法

ここにいう右隅統語解析法とは、文法規則の右辺の最後から始め、左端に向けて解析を進めていくような解析法の総称である。したがって、必ずしも、いわゆる左隅統語解析法^[10]と対立する概念というわけではない。しかしながら、左隅統語解析法は、文法規則の右辺の先頭から始め、右端に向けて解析を進めていく解析法であるので、この意味では対立していると考えられることもできる。

3. 1. 左隅統語解析法と右隅統語解析法

次のような文法規則(式1)があったとしよう。

$$A \rightarrow B_1 B_2 \dots B_n \quad (\text{式1})$$

このとき、左隅統語解析法は、 B_1 から始めて順に B_n まで処理を進める。これに対して、右隅統語解析法は、逆に、 B_n から始めて順に B_1 まで処理を進めるのである。左隅統語解析法も右隅統語解析法もともに、上昇型統語解析法^[10]の一種である。ただし、右隅統語解析法に関しては、これ以上の制約は全くない。つまり、後戻り型^[10]の右隅統語解析法を考へることもできれば、並行型^[10]の右隅統語解析法を考へることもできるのである。

3. 2. 右隅統語解析法のアルゴリズム

次に、並行型の右隅統語解析法に対するアルゴリズムを示す。これは、次のような二種類の手続きからなる。

[手続き1]

不活性弧 e_i が、節点 u と v の間を張っているものとする。また、 e_i の文法範疇を a とし、対応する項を a' とする。このとき、 a を右辺の末尾に持つすべての文法規則(式2)に対して、 $[[?]p_1 \dots [?]p_n a']p$ のような項を持つ活性弧を u と v の間に追加せよ。

$$p \rightarrow p_1 \dots p_n a \quad (\text{式2})$$

[手続き2]

不活性弧 e_i が節点 u と v の間を張り、活性弧 e_a が節点 v と w の間を張っているものとする。つまり、二つの弧は節点 v を介して隣合っているものとする。このとき、 e_a の

項の最後の空所が $[?]a$ であり、 e_i の文法範疇が a であるならば、 e_a の項の最後の空所を e_i の項で置き換えたものを項とするような弧を u と w の間に張れ。

なお、図3～5は、次のような文に対する解析例である。

(文2) 虎が草原を走る

3. 3. 半活性弧

図2の文法規則に、次のようなものがある。

- | | |
|--------------------------|--|
| 1. $s \rightarrow ppv$ | 4. $v \rightarrow \text{走る}$ |
| 2. $s \rightarrow ppppv$ | 5. $n \rightarrow \text{虎} \mid \text{草原}$ |
| 3. $pp \rightarrow np$ | 6. $p \rightarrow \text{が} \mid \text{を}$ |

図3. 文法規則の例(その2)

#	項
1	$[\text{虎}]_n$
2	$[\text{が}]_p$
3	$[\text{草原}]_n$
4	$[\text{を}]_p$
5	$[\text{走る}]_v$
6	$[[?]_n [\text{が}]_p]_{pp}$
7	$[[\text{虎}]_n [\text{が}]_p]_{pp}$
8	$[[?]_n [\text{を}]_p]_{pp}$
9	$[[\text{草原}]_n [\text{を}]_p]_{pp}$
10	$[[?]_{pp} [\text{走る}]_v]_s$
11	$[[[\text{草原}]_n [\text{を}]_p]_{pp} [\text{走る}]_v]_s$
12	$[[?]_{pp} [?]_{pp} [\text{走る}]_v]_s$
13	$[[?]_{pp} [[\text{草原}]_n [\text{を}]_p]_{pp} [\text{走る}]_v]_s$
14	$[[[\text{虎}]_n [\text{が}]_p]_{pp} [[\text{草原}]_n [\text{を}]_p]_{pp} [\text{走る}]_v]_s$

図4. 右隅統語解析によるチャート法

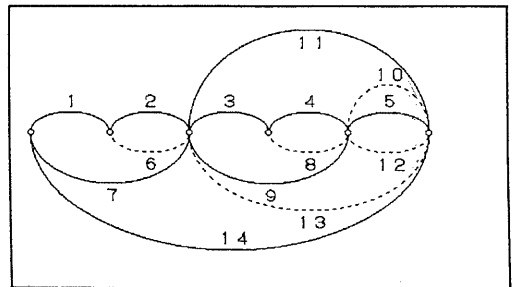


図5. 右隅解析によるチャートのグラフ表現

(規則 4) $ap \rightarrow a$
 $ap \rightarrow adv\ ap$

そして、 ap が左辺に出現する規則はこれだけである。したがって、これらの規則は、下の(規則 5)のように書いてもまったく同じことである。

(規則 5) $ap \rightarrow adv^* a$

ただし、 x^* は範疇 x を 0 回以上繰り返してもよいことを表している。そして、これはさらに、(規則 6)のように書いても本質は変わらない。

(規則 6) $a \rightarrow adv^* a$

このような考え方に従って、図 2 の文法規則を書き換えたものを図 6 に示す。

$a \rightarrow adv^* a$	$v \rightarrow v\ auxv^*$
$n \rightarrow (a v)^* n$	$v \rightarrow pp^* v$
$pp \rightarrow n\ p$	

図 6. 文法規則の例 (その 3)

さて、(規則 6) のようなものは活性弧と不活性弧を使って処理するより、半活性弧というものを使って処理した方が、解析に必要な弧の数が少なく済む。半活性弧に対応する項を、たとえば、次の(項 5)のように表すことにする。

(項 5) $[[([?]_{adv})^* [細やかな]_a]_a$

これは、文法範疇が adv である不活性弧に対応する項を 0 個以上いくつでも取り込めることを表現している。また、他の活性弧に対応する項は、半活性弧に対応する項を取り込むこともできるし、半活性弧に対応する項自身、別の半活性弧に対応する項を取り込んでもよい。このように、半活性弧は、活性弧と不活性弧の中間の性質を持っている。

4. 消去可能性

チャート法を用いて統語解析を行なう場合、一度作られた弧は活性弧であれ不活性弧であれ、決して消去されるようなことはない。それは、少なくとも一つの文に対する統語解析が終了するまでは存在し続ける。ところが、

生成された弧のうち不要になったものを早めに消去すれば、これによって無駄な部分解析木の生成を防ぐことができる。そして、これは、無駄な部分解析木を削除するための統一的な方法となりうる。つまり、到達可能性や接続可能性で排除していた部分解析木だけでなく、もっと大域的な制約の必要な部分解析木も削除することができる。

4. 1. 半活性弧の消去

もう一度、右隅統語解析法のアルゴリズムに戻ってみよう。半活性弧も使うことにすると、このアルゴリズムの手続き 2 は、活性弧(半活性弧)と不活性弧(半活性弧)を結びつけて、新しい弧を生成するようなものになる。特に、半活性弧に不活性弧(半活性弧)を結びつけるようなものを考えよう。この結果生成されるのは、再び半活性弧である。さて、半活性弧に不活性弧(半活性弧)を結びつける場合、もともとは右側にあった半活性弧を消去できる場合がある。これを、ここでは、消去可能性と呼んでいる。

なお、これから先、弧のことをオブジェクトと呼ぶことにする。それは、消去動作において、弧は単なるデータというよりは自律的なオブジェクトと見なせるからである。オブジェクトはその内部に、対応する文法範疇や項を持っているものとする。

4. 2. 消去定理

現在分かっているのは、消去可能性を保証する十分条件だけである。以下、統語解析に右隅統語解析法を使用するとして、この十分条件について述べる。まずは、簡単な定義を行なう。

[定義 5]

文法範疇 p に対して、 p から文法規則によって導かれる記号列の中に、範疇 q を含むものが一つでもある場合に、 p は q を包含可能であるという。

この定義を用いて、次のような第一種消去定理を記述できる。ただし、証明は省略する。

[定理 1] 第一種消去定理

文法カテゴリ a を右辺に持つ規則が、次のような(式 3)だけであるとすると。

$$p \rightarrow (p_1 | \dots | p_n)^* p \quad \dots \quad (式 3)$$

だけであるとする。ただし、 p は語彙範疇でもあるとし、 $a = p_1$ である。このとき、文法範疇 $p_1 \dots p_n$ が、語彙範疇 p を包含可能でないならば、 p という範疇の半活性オブジェクトが、 a に対応するオブジェクトと結合するときに、もとの半活性オブジェクトを消去することができる。

さて、図 6 の文法規則から、形容詞 (a) に対応する半活性オブジェクトが副詞 (adv) に対応するオブジェクトと結合するとき、もとの半活性オブジェクトを消去可能であることが分かる。なぜなら、 adv を右辺に持つ文法規則は、次の (規則 7) に限られ、さらに、 adv は語彙範疇なので語彙範疇 a を包含可能でないからである。

$$(規則 7) \quad a \rightarrow adv^* a$$

そこで、もう一度、(文 1) を解析することを考えてみよう。「細やかな」という形容詞 (a) が、「実に」という副詞 (adv) と結合するとき、新たに「実に細やかな」という形容詞が生成されるが、このとき、もとの「細やかな」という形容詞は消去可能であった。したがって、(項 2) のように「細やかな」という形容詞が「神経」という名詞と結合することはなく、無駄な解析木を抑制することができる。ここでは、消去可能性は到達可能性と同じ働きをしている。

余白の都合で省略するが、定理 1 に類似した別の消去定理を用いると、動詞 (v) と助動詞 (adv) が結合するとき、もとの動詞を消去可能なが分かる。ただし、この場合は例外的に通常のチャート法に従って解析を行なうものとする。そうすると、(文 1) の解析において、「持つ」という動詞 (v) と「ている」という助動詞 (adv) が結合するとき「持っている」という動詞が新たに生成され、もとの「持つ」という動詞は消去されることになる。したがって、(項 3) のようなオブジェクトが作られることはなく、これは、連接可能性を利用するのと同じことになる。

4. 3. 大域的制約を用いる消去定理

次に述べる第二種の消去定理は、大域的制約を用いるものである。

[定理 2] 第二種消去定理

文法カテゴリ a を右辺に持つ規則が、

$$p \rightarrow (p_1 | \dots | p_n)^* p \quad \dots \quad (式 4)$$

だけであるとする。ただし、 p は語彙範疇でもあるとし、 $a = p_1$ である。このとき、文法範疇 p の半活性オブジェクトの右側に、語彙範疇 p に対応するオブジェクトが存在しないならば、この半活性オブジェクトが、 a に対応するオブジェクトと結合するときに、もとの半活性オブジェクトを消去することができる。

この定理を用いると、(文 1) の解析において、「神経」という名詞 (n) と、「実に細やかな」という形容詞 (a) が結合して、「実に細やかな神経」という名詞ができるとき、もとの「神経」という名詞は消去可能であることが分かる。なぜなら、 a を右辺 (末尾を除く) に持つ規則は、次の (規則 8) だけであり、さらに、「神経」という名詞の右側には、別の名詞は存在しないからである。

$$(規則 8) \quad n \rightarrow (a|s)^* n$$

そして、「神経」という名詞が消去されるのであれば、(項 4) のような「神経」を名詞として含むような部分解析木が作成されないのは明らかである。前にも述べたように、このような部分解析木は、到達可能性や連接可能性を利用して削除できず、消去可能性を用いて初めて削除することができる。

5. 統語解析実験

我々は、①通常のチャート法に何も適用しないもの、②通常のチャート法に到達可能性を適用したもの、③通常のチャート法に連接可能性を適用したもの、④右隣統語解析法に消去可能性を適用したものの 4 種類について統語解析実験を行なってみた。ただし、消去可能性については、上で述べた消去定理に対して多少の拡張を行なっている。実験は、日本語の文法規則を用いて行なった。規則数は語彙規則を除いて 15 個であるが、規則数が比較的少ないのは、HPSG に類似した単一化文法を用いているからである。例文は、朝日新聞社の天声人語の中から、適宜抽出した。5 個の文に対して行なった実験結果を表 3 に示す。

番号	無制約	到達可能性	接続可能性	消去可能性
①	116 (49%)	116 (49%)	93 (61%)	18 (100%)
②	428 (41%)	428 (41%)	291 (61%)	73 (99%)
③	191 (51%)	191 (51%)	146 (66%)	36 (100%)
④	1148 (32%)	1148 (32%)	749 (49%)	199 (100%)
⑤	102 (43%)	90 (49%)	76 (58%)	13 (100%)

表3. 生成される弧の総数
(括弧内は弧の利用率)

数値は、解析時に生成された弧（オブジェクトあるいは部分解析木）の総数であり、括弧内に、弧の利用率を示している。利用率というのは、生成された弧の全体に対する無駄でない弧の割合である。これを見て分かるように、チャート法に何も適用しないときは、半数以上の弧が無駄なものである。そして、到達可能性を適用しても、無駄な弧はほとんど削減されていない。接続可能性を利用すると無駄な弧をある程度まで排除することができるが、それでも依然としてかなりの量の無駄な弧が残っていることも分かる。これに対し、右隅統語解析法に消去可能性を適用した場合、無駄な弧はほとんど生成されていない。また、半活性弧を利用しているため、必要な弧の絶対量が減少していることも見てとれる。

なお、多くの統語解析アルゴリズムで採用されている到達可能性は、少なくとも我々が用いた文法規則に対しては有効でないことが明らかになった。原因として、日本語文法の特殊性や文法規則の少なさを挙げることもできるかも知れない。

6. まとめ

文脈自由文法を用いた統語解析において無駄な部分解析木を削減する方法として、右隅統語解析法に消去可能性を適用するやり方を提案した。消去可能性は、従来の到達可能性や接続可能性とは異なり、大域的な制約を用いている。そして、この方法で無駄な部分解

析木をほぼ完全に削除できることを確認した。

今後の課題として残されているのは、消去定理の一般化である。現在の消去定理は、消去可能であるためのかなり強い十分条件を述べているだけで、もっと一般の場合を扱っていない。そこで、別の観点からこれを見直し、必要十分条件にさらに近づける必要があると思われる。

《参考文献》

- [1] Shieber, S.M., et al.: The Formalism and Implementation of PAT R-II, Research on Interactive Acquisition and Use of Knowledge, SRI International (1983).
- [2] Gunji, T.: Japanese Phrase Structure Grammar: A Unification-based Approach, Dordrecht, D. Reidel (1987).
- [3] 大澤, 米澤: オブジェクト指向方式による対話理解システム, コンピュータソフトウェア, Vol.2, No.1, pp.11-28 (1985).
- [4] 佐藤, 田淵, 長尾: CFGパーサのオブジェクト指向言語による実現法とその並列化, WOOO-87 (1987).
- [5] Pratt, V.R.: The LINGOL System, MIT AI (1975).
- [6] Matumoto, Y. et al.: BUP: A Bottom-Up Parser Embedded in Prolog, New Generation Computing, Vol.1, No.2 (1983).
- [7] 松本, 杉村: 論理型言語に基づく構文解析システム SAX, コンピュータソフトウェア, Vol.3, No.4 (1986).
- [8] 上原, 豊田: 先読みと予測機能を持つ述語論理型構文解析プログラム: PAMP S, 情報処理学会論文誌, Vol.24, No.4, pp.496-504 (1983).
- [9] Kay, M.: Algorithm Schemata and Data Structures in Syntactic Processing, Technical Report CSL-80-12, Xerox PARC, Oct. (1980).
- [10] 松本: 第3章統語解析の手法, 自然言語理解, オーム社 (1988).