

多元的類似度計算に基づく文脈を考慮したボトムアップ構文解析法

側嶋康博

ATR音声翻訳通信研究所

sobasima@itl.atr.co.jp

用例との類似度計算のみでボトムアップに自然言語を構文解析する手法を提案する。本手法の特徴は、局所文脈を含む多元的な類似度の導入により、高い精度のまま頑強な構文解析が可能となることである。すなわち、意味的に近い用例がない場合は構文的に近い用例を模倣し、構文・意味的に類似する用例が複数ある場合は、局所文脈の類似度を用いて文脈を考慮した解の選択を行うことができる。また、各種の類似度解析(構文・意味・局所文脈)を統合して実行するため処理効率が高い。実験により、用例追加による副作用が極めて小さいことが示されている。

A Context-dependent, Bottom-up Parsing Method Based on Multi-dimensional Similarity Calculations with Partial Tree Examples

Yasuhiro Sobashima

ATR Interpreting Telecommunications Research Laboratories

This paper proposes a new method for parsing natural language expressions: Multi-dimensional Analogy-based Context-dependent Bottom-up Parsing (MAC-BP). This method performs accurate and robust parsing by utilizing multi-dimensional similarity calculations, including local context similarity, with partial tree data extracted from actual examples. In this paper an outline of a linguistic model introduced for the similarity-based analysis will be presented as well as the similarity calculations for the bottom-up parsing, and the experimental results from a MAC-BP prototype system which is accurate, robust, and efficient in parsing spoken dialogues.

1 はじめに

自然言語の構文解析は、自然言語処理の中で最も基本的かつ重要な技術であるが、自然言語が持つ自由度(語彙的・構造的曖昧性、語句の省略、非文法的表現など)のため、高い精度で頑強に構文解析を行うことは依然困難である。従来のルールに基づく手法では、記述を詳細化して精度を上げようとすると頑強さを失い、逆に記述を粗くすると多くの解釈から正解の選択が困難になるため、数多くのヒューリスティック・ルールを駆使して副作用に対処しながら複雑な制御を行わなければならなかった。

近年、用例との類似性に着目したアプローチが提唱^[1]され、主に機械翻訳の変換^{[2],[3],[4]}に応用されて成果を収めている。このアプローチでは、語のソーラス階層関係から得られた、用例との意味的な類似度または距離^[3]を基に、尤もらしい変換事例を解として採用する。この手法では、アナログ的な数値を判断に使用して頑強さを向上させ、また具体的な用例を多く用意することにより精度を向上させる。すなわち、高い精度と頑強さを共に備えた手法として期待される。しかしながら、これまでの用例に基づくアプローチでは、類似度は「語の意味的類

似度」に、適用対象はルールで曖昧性解消できない「変換パターンの選択」に限定されており、類似度尺度だけで自然言語を構文解析することは行われていなかった^(注1)。

本稿では、部分木用例データとの多元的な類似度計算により、類似度尺度のみでボトムアップに自然言語を構文解析する手法を提案する。本手法の特徴は、局所文脈を含む多元的な類似度の導入により、高い精度のまま、頑強に構文解析することである。すなわち、意味的に近い用例がない場合でも機能的(構文的)に近い用例を模倣し、また機能・意味的に類似する用例が複数ある場合は局所文脈の類似度を用いて文脈を考慮した解の選択を行うことができる。また、種々の解析(機能・意味・形態および局所文脈)を統合して実行するため、処理効率が高い。例えば、ボトムアップな解析の各段階でトップダウン制約(局所文脈の制約)がかかるため、全探索やバックトラックなしに効率的に構文解析する。さらに、実験を通して、用例追加による副作用が極めて小さいことが示されている。

注1 TDMT^[4]では、解釈可能な構造はすべて求めた上で、用例と意味的に最も近い解を選択する。

2 用例を用いた構文解析の基本アイデア

本研究の基本的立場である用例に基づくアプローチの利点の1つは、人間の経験的判断に近い処理を行うことである。人間が自然言語を認識・理解する場合、図1のように、基本的には、知識と照合してボトムアップに意味的なまとまりを作りながら全体を認識・理解していると推察される。

この過程をシミュレートする構文解析では、語彙(辞書)と句構成(部分木用例)の知識が必要であり、パーサは、それら知識をアクセスしながら類似度計算を行い、尤もらしい意味的なまとまりを判断してノードを作成する。処理は、解析の終了、すなわち全体が1つのノードにまとまるまで繰り返す。ただし、曖昧性解消の必要^(注2)から、部分木用例に局所文脈情報も加えておく。

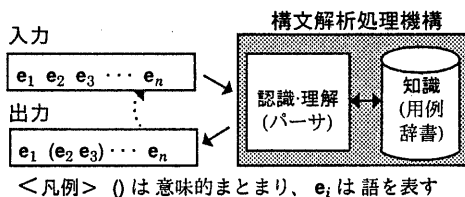


図1 用例ベースの構文解析処理

3 類似度解析のための言語モデル

3.1 解析の単位

本研究では、話し言葉の発話が対象であり、翻訳を応用例に考えている。そのため日英表現の対応の便宜^[5]も考慮して、句点で分けられる「文」に限定しない解析単位を設定する。すなわち、応答表現(「はい」「yes/okay」など)から末尾の働きかけ表現(「～さん」「right?/please」など)までを構文解析の単位(メッセージ)としている。トップノードであるメッセージは、図2に示すように、プレ、メイ

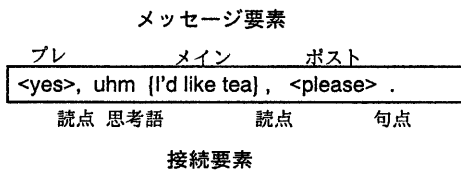


図2 メッセージの構成例

注2 例えば「... make a room reservation」という入力に対して「a+room→a room」「room+reservation→room reservation」という2つの部分木用例の類似度だけでは選択できないが、左右語句(左端・右端マーカを含む)情報の追加により曖昧性が解消できる。

ン(必須)、ポストの各メッセージ要素と、それらを接続する要素(ジョイント)からなる。

3.2 発話の言語的構成

プレ、メイン、ポストの各メッセージ要素は、語または複数の語で構成される句からなる。以下では語を要素、句を構造と呼ぶ。

要素は「自立語的要素」と、他の自立語的要素と結合して構造を作る「付属語的要素」とに分類される。付属語的要素は、日本語では助詞や助動詞が対応し、英語では、冠詞、前置詞、助動詞などが対応する。したがって、本言語モデルにおいて、発話の言語的構成は図3のようになる。

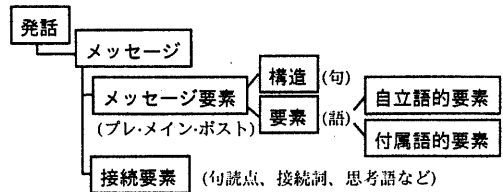


図3 発話の言語的構成

本モデルでは、統一した類似度計算の都合上、自立語的構造だけを構造と認める。例えば、日本語の後置詞句、英語の前置詞句は構造と扱わない^(注3)。

3.3 役割を用いた構造の記述

ある構造とその構成要素との関係を役割と呼び、ヘッド、サブ、フラグメントの3通りを設定する。ヘッドは構造の意味的中心、サブはその補語または付加語、フラグメントは構造を構成するために必須な付属語的要素である。また、ヘッドとサブは自立語的要素または構造である。さらにヘッドは、その構造に活用の制約を受けないストロング(S)ヘッドと、制約を受けるウィーク(W)ヘッドの2通りに分ける。図4に英文の構造と役割の例を示す。

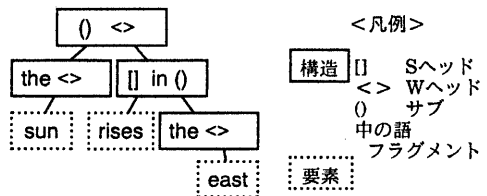


図4 役割を用いたメイン・メッセージ要素「the sun rises in the east」の構造記述例

注3 断片的な発話で、本来の係り先が省略された前置詞句や後置詞句(「just for myself」「お名前は」など)は構造として扱う。

4 多元的類似度計算

4.1 要素間類似度

ある言語的要素 A は、いくつかの属性に分解できるものとする。例えば、機能(品詞・活用などの構文的属性)、意味、形態の3つの属性 $A_{fun}, A_{sem}, A_{gra}$ のセットとして eq. 1 のように表す。

$$A(A_{fun}, A_{sem}, A_{gra}) \quad (\text{eq. 1})$$

A と他の要素 B との要素間類似度 $E\text{-Sim}$ を eq. 2 で定義する(注4)。

$$E\text{-Sim}(A, B) = \phi(\sigma_{fun}, \sigma_{sem}, \sigma_{gra}) \quad (\text{eq. 2})$$

ここで、 ϕ は要素間類似度を与える関数であり、 $\sigma_{fun}, \sigma_{sem}, \sigma_{gra}$ はそれぞれ、 A, B 間の機能、意味、形態の類似度である。

4.2 構造間類似度

複数の要素 $A_1, A_2, \dots, A_n$ からなる列 A が構造を作る場合、同数の要素からなる列 B との間の構造間類似度を eq. 3 で定義する。

$$S\text{-Sim}(A, B) = (1/W) \sum_i E\text{-Sim}(A_i, B_i) w_i \quad (\text{eq. 3})$$

ここで、 i は要素位置を表す添字であり w_i は役割に応じた重みである。また、正規化係数 $(1/W)$ と w_i との間に $W = \sum_i w_i$ の関係がある。

4.3 文脈的類似度

図5に示すように、言語的表現 A, B が左右にそれぞれ a_1, a_2 と b_1, b_2 の語句を伴っている(注5)とする。

$$\dots a_1 A a_2 \dots, \dots b_1 B b_2 \dots$$

図5 2つの言語的表現 A, B の文脈的關係

この場合の A, B 間の文脈的類似度(局所文脈の類似度) $C\text{-Sim}$ を、文脈的類似度を与える関数 ϕ_c を用いて eq. 4 で定義する。

$$C\text{-Sim}(A, B) = \phi_c(E\text{-Sim}(a_1, b_1), E\text{-Sim}(a_2, b_2)) \quad (\text{eq. 4})$$

4.4 統合した類似度

構文解析で解選択の尤度として使用する、構造 A, B 間の統合した類似度 Sim を、構造間類似度と文脈的類似度の関数 Φ である (eq. 5) と定義する。

$$\text{Sim}(A, B) = \Phi(S\text{-Sim}(A, B), C\text{-Sim}(A, B)) \quad (\text{eq. 5})$$

注4 付属語的要素間の要素間類似度は、一致した場合1、不一致の場合0と定義する。

注5 A, B が左端にある場合 a_1, b_1 は左端マーカ、右端にある場合 a_2, b_2 は右端マーカを使用する。

5 類似度計算を用いたボトムアップ構文解析

5.1 入力と部分木用例データ

図6に示すように、構文解析の入力は、属性セットからなる要素の列とする。また、図7に示す文脈付き部分木用例データを用意する。

$$e_1(e_{1f}, e_{1s}, e_{1g}) \quad e_2(e_{2f}, e_{2s}, e_{2g}) \quad \dots \quad e_n(e_{nf}, e_{ns}, e_{ng})$$

図6 入力する要素列

部分木用例データ " $N | R | C_l | e_1 | e_2 | \dots | C_r$ "

N 構造を作る要素の数
 R 構造全体の属性(機能・意味・形態)セット
 e_i i 番目の要素の属性(機能・意味・形態)セット
 C_l 左に隣接する要素の属性セット(左文脈要素)
 C_r 右に隣接する要素の属性セット(右文脈要素)

図7 文脈付き部分木用例データ

5.2 類似度計算によるノードの作成

構文解析では、連続する入力要素列の組合せを変えながら、部分木用例データベースから検索した個々の用例との間で類似度計算を行い、統合した類似度が最大となる要素列を選び用例の属性を模倣して新ノードを作成する(図8参照)。同点の他の要素列については優先条件(後述)を適用し、また新ノード属性候補が複数ある場合は、保存する。この処理をノード数(要素数)が1になるまで繰り返す。

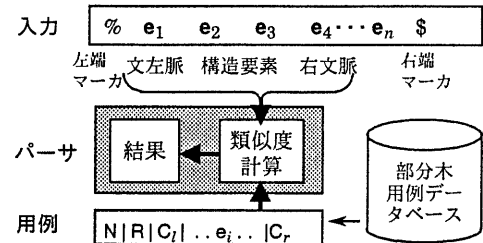


図8 文脈付き用例を用いた構文解析

5.3 曖昧性の解消

1要素に複数属性セットがある場合、最大の統合した類似度を与える属性セットを選択することにより、曖昧性を解消する。すなわち、構造的曖昧性を解消しながら語彙的曖昧性を解消する。図9は、「三日」という表現の意味属性が係り先(ヘッド)により選択される例である。

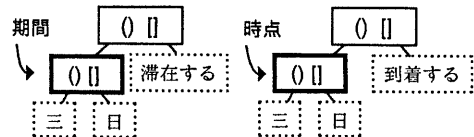


図9 語彙的曖昧性の解消例

6 実験

6.1 用例データ

実験には、表1の用例データを使用した。

表1 実験に用いた用例の種類と量 (M:メッセージ)

種類 (言語)	M数	文数	語数	語数/M
動詞型別 (日)	259	259	2639	10.2
基本文 ^[6] (英)	259	259	1930	7.5
会話 (日)	835	852	6211	7.4
基本文 ^{[7],[8]} (英)	835	853	5527	6.6
ATR旅行 (日)	545	666	6997	12.8
模擬会話 ^[9] (英)	545	660	6599	12.1
合計 ^(注6) (日)	1639	1777	15839	9.7
(英)	1639	1772	14056	8.6

6.2 類似度の設定

6.2.1 機能属性と機能類似度

機能属性は、大中小3階層に分類した品詞と、各品詞大分類ごとに共通な活用とを定めた。表2に品詞大分類、表3に英語の名詞類^(注7)の分類を示す。品詞分類の総数は、日本語 86、英語 90 である。

表2 品詞大分類

日本語品詞	英語品詞
N 名詞類	N 名詞類
Q 数詞類	Q 数詞類
A 形容詞類	A 形容詞類
V 動詞類	V 動詞類
	B be 動詞句
	X 助動詞句
D 副詞類	D 副詞類
S 文	S 文
R 間投表現	R 間投表現

表3 英語名詞類細分

N10 名詞
N11 可算名詞
N12 不可算名詞
N20 名詞句
N30 代名詞
N31 主格代名詞
N32 目的格 "
N33 再帰目的格 "
N34 指示 "
N35 否定 "
N40 固有名詞
N50 単位
N51 数量単位
N52 位置単位

品詞分類、活用それぞれに3桁、1桁のコードを与え、それらを使用して、品詞類似度 σ_{pos} と活用類似度 σ_{inf} を eq. 6, eq. 7 により求める。

$$\sigma_{pos} = n/3 \quad (n = 0, 1, 2, 3) \quad (\text{eq. 6})$$

$$\sigma_{inf} = 0, 1 \quad (\text{eq. 7})$$

ここで、 n は品詞コードが上位から一致した桁数であり、 σ_{inf} はコードが一致した場合 1、不一致の場合 0 を与える。

また、eq. 8 に示すように、機能類似度 σ_{fun} は品詞類似度 σ_{pos} と活用類似度 σ_{inf} の関数とする。

$$\sigma_{fun} = \phi_{fun}(\sigma_{pos}, \sigma_{inf}) \quad (\text{eq. 8})$$

注6 部分木用例総数は日本語8887、英語7591である。

注7 代名詞の格変化は品詞細分として扱う。

6.2.2 意味属性と意味類似度

意味属性の分類は、日英両言語で共通とし、語句レベルの意味分類とメッセージレベルの発話行為タイプの分類^[6]の2通り、それぞれ大中小3階層の分類(3桁コード)とした(表4を参照)。細分化された意味属性の分類総数は323である。また、意味類似度 σ_{sem} についても、意味分類コードの上位からの一致桁数 n を用いて eq. 9 で与える。

表4 意味属性の分類

語句	発話行為タイプ ^(注8)
1 人	- 反応
2 物	-1 反応的応答(驚き・確認・理解)
3 抽象	-2 陳述的応答(判断・情報・約束)
4 時	0 思考
5 所	01 独り言(理解・判断・情報)
6 量	02 思考語(内容思考・表現思考・語思考)
7 様態	+ 作用
8 出来事	+1 情報提供(事象・心情・判断)
9 行為	+2 要求(判断・情報・行為)

$$\sigma_{sem} = n/3 \quad (n = 0, 1, 2, 3) \quad (\text{eq. 9})$$

6.2.3 形態属性と形態類似度

形態属性は、対象が要素の場合、見出し語表記を採用する。構造の場合は、語句構成の利用も考え、フラグメントの文字列と共にSヘッド、Wヘッド、およびサブの各要素の記号を加えて表す。ただし、抽象度を高めるため、サブ要素の表記は除外する。図10に形態属性の例^(注9)を示す。

(日本文例) "部屋の予約をしたいのです"
形態属性 "<<()>[し]>たい>ののです"
(英文例) "I would like to make a room reservation"
形態属性 "<()>would <[like] to ()>"

図10 文の形態属性例

形態類似度は、形態属性の文字列、形態属性中のヘッド文字列および基底パターン(ヘッド・サブ記号およびフラグメントの文字列^(注10))の一致を用いて、eq. 10 により与える。

$$\sigma_{gra} = \begin{cases} 1 & (\text{形態属性が完全に一致}) \\ 0.9 & (\text{ヘッド・基底パターンが一致}) \\ 0.6 & (\text{ヘッドのみ一致}) \\ 0.3 & (\text{基底パターンのみ一致}) \\ 0 & (\text{上記以外}) \end{cases} \quad (\text{eq. 10})$$

注8 発話行為タイプの分類の括弧内は小分類を示す。

注9 日本語の動詞は、名詞派生語との関連を考慮し連用形(ここでは「し」)を見出しに使用している。

注10 図10の例では「<>ののです」「() would <>」が基底パターンとなる。

6.2.4 類似度関数・係数および重み

eq. 2 の要素間類似度は、各属性の類似度ごとの関数(類似度関数)の積(注11)で表わし(eq. 11)、各類似度関数 ϕ_x は、それぞれの類似度 σ_x の一次関数とする(eq. 12)。ここで、 μ_x (類似度係数, $x = \text{pos, inf, sem, gra}$) は要素間類似度関数 ϕ における各属性ごとの類似度 σ_x の重みを表し(注12)、実験的に求める(6.5を参照)。

$$\begin{aligned} \phi(\sigma_{\text{fun}}, \sigma_{\text{sem}}, \sigma_{\text{gra}}) \\ = \prod_x \phi_x(\sigma_x) \quad (x = \text{pos, inf, sem, gra}) \quad (\text{eq. 11}) \\ \phi_x(\sigma_x) = \mu_x \sigma_x - \mu_x + 1 \quad (\text{同上}) \quad (\text{eq. 12}) \end{aligned}$$

eq. 3 の構造間類似度 S-Sim の計算に必要な役割ごとの重みを eq. 13 で与える。

$$w = \begin{cases} 10 & (\text{ヘッド}) \\ 5 & (\text{メイン・メッセージ要素}) \\ 1 & (\text{サブ、プレ・ポスト各メッセージ要素}) \end{cases} \quad (\text{eq. 13})$$

eq. 4 の文脈的類似度(C-Sim または σ_{con}) は、左右要素の類似度の積で与える(eq. 14)。ここで σ^l, σ^r はそれぞれ左・右要素の要素間類似度である。また、統合した類似度関数 Φ は、文脈類似度関数 ϕ_{con} および係数 μ_{con} を用いて eq. 15, eq. 16 で表す。

$$\phi_c = \sigma_{\text{con}} = \phi(\sigma^l) \phi(\sigma^r) \quad (\text{eq. 14})$$

$$\Phi = \text{S-Sim} * \phi_{\text{con}}(\sigma_{\text{con}}) \quad (\text{eq. 15})$$

$$\phi_{\text{con}}(\sigma_{\text{con}}) = \mu_{\text{con}} \sigma_{\text{con}} - \mu_{\text{con}} + 1 \quad (\text{eq. 16})$$

6.3 新ノード属性

ある要素列(ヘッド H_{in}) に与える新ノード(N) の品詞・活用・意味属性は、採用された部分木用例データの構造ノード(E)とそのヘッド(H_{eg})の属性を使って、eq. 17, eq. 18 のように模倣する。また、形態属性は、6.2.3 の手順で作成する。

$$N_{\text{pos}} = E_{\text{pos}} \quad (\text{eq. 17})$$

$x = \text{inf, sem}$ の場合

$$N_x = \begin{cases} H_{\text{in}_x} & (H_{\text{eg}_x} = E_x) \\ E_x & (H_{\text{eg}_x} \neq E_x) \end{cases} \quad (\text{eq. 18})$$

6.4 優先条件

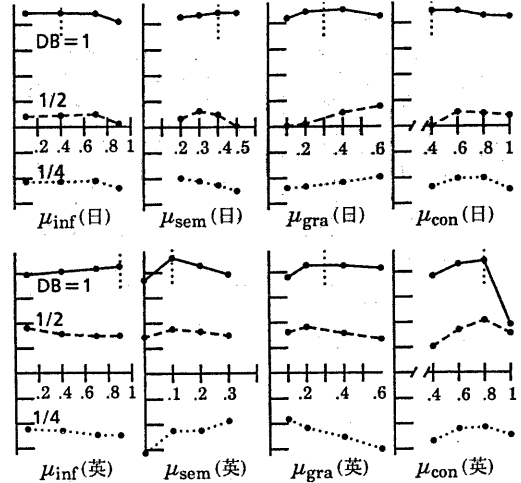
新ノードとなる要素列の決定は、(1) 統合した類似度の高いもの、もし複数の同点があれば(2) 要素数が多いもの、もし同数であれば(3) 日本語は左端寄り、英語は右端寄り、の順に優先する。

注11 eq. 8 の機能類似度 ϕ_{fun} は、品詞分類の類似度 $\phi_{\text{pos}}(\sigma_{\text{pos}})$ と活用の類似度 $\phi_{\text{inf}}(\sigma_{\text{inf}})$ の積となる。

注12 活用の制約を受けないSヘッドについては、活用の一致は無視する($\mu_{\text{inf}} = 0$)。

6.5 類似度係数決定実験

この実験では、用例からランダムに選んだ日英被験文各80メッセージについて、品詞の類似度係数($\mu_{\text{pos}} = 1$)以外の係数値を組合せ(注14)被験文の部分木データは学習させず、構文解析実験を行った。図11に、各係数ごとの成功得点(注13)のグラフを示す。この結果から、表5の係数値を採用する。



<凡例> 上段: 日本語, 下段: 英語 (DBサイズは全体を1)
縦軸: 得点合計 (目盛:300点), 横軸: 各係数値

図11 類似度係数決定実験結果

表5 採用した類似度係数値

言語	μ_{pos}	μ_{inf}	μ_{sem}	μ_{gra}	μ_{con}
日本語	1	0.4	0.4	0.3	0.4
英語	1	0.9	0.1	0.3	0.8

6.6 クローズテスト

準備した用例からランダムに被験文100メッセージを選び、このデータを含めたまま学習データの量を半減させて構文解析を行った。図12はこの実験を3回行って得た結果の平均を表すグラフである。また、図13に、同じ実験で文脈類似度を無視した場合($\mu_{\text{con}} = 0$ と設定)の結果を示す。

6.7 オープンテスト

クローズテストと同様の手順(100メッセージ3セット)でオープンテストを行った。ただし、このテストでは、被験文の部分木データは学習データに含めていない。図14にその結果を示す。

注13 構造と属性の正解に2点、構造のみ正解に1点、誤った構造に-1点、解析失敗を-2点とした。

注14 係数 ($4 \times 4 \times 4 \times 4 = 256$), 被験文 (80), 言語 (2), 学習データサイズ (3) を組合せ、計122,880 回行った。

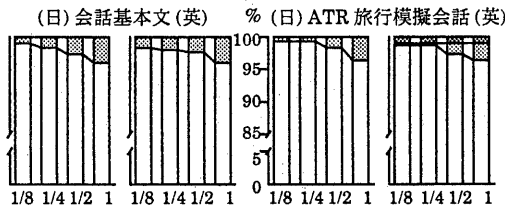


図12 クローズテスト結果

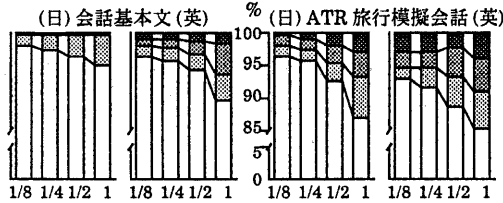


図13 クローズテスト ($\mu_{con}=0$) 結果

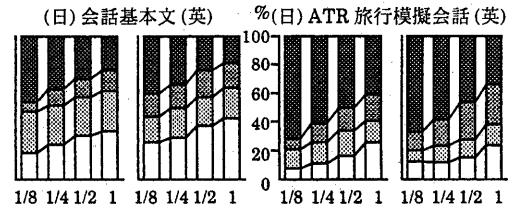


図14 オープンテスト結果

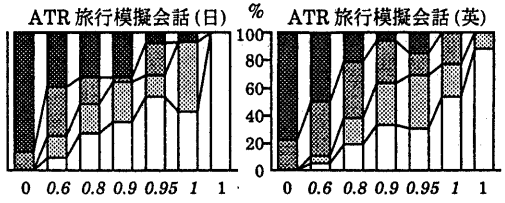


図15 クリティカル類似度とオープンテスト結果

<凡例> 縦軸 結果の割合(%) □ 構造・属性正解 ■ 構造正解 ■ 誤り ■ 失敗
横軸 図12~14:学習データ量(全体を1), 図15:クリティカル類似度(斜体はその値未満を表す)

図15は、パース対象メッセージの構成部分木のうち、用例との類似度(用例中の全部分木と比較した場合の最大値)が最低となる部分木の類似度(クリティカル類似度と呼ぶ)とパース結果を表すグラフである。このグラフから、クリティカル類似度に比例してパース精度が向上しており、類似度設定の妥当性を示している。

6.8 結果の考察

今回の実験では、メッセージ隣接用例を用いた属性の曖昧性解消は行っていない(注15)ため、模範解答と属性が異なり、「構造正解」と判定されたものが多くなっているが、クローズテストでは、文脈類似度を利用することにより、用例追加による副作用を抑え、96%以上の精度を保持している。また、オープンテストでは、用例の増加に比例して精度が向上しており、本手法の有効性を示している。

7. まとめと課題

隣接語句の文脈を含む部分木用例との多様な類似度計算に基づくボトムアップな構文解析手法の概要と、いくつかの実験結果を示した。本手法は、構文・意味・用法的制約を機能・意味・形態および局所文脈の類似度に置き換えることにより、制御を容易にし、かつ処理の頑強さを向上させている。特に、局所文脈の利用により、用例知識の追加による副作用を抑え、インクリメンタルな性能向上が可能である

注15 例えば、「そうですね」という表現の発話行為タイプ(反応的応答か判断要求か)の判定に直前メッセージ情報を使用していない。

ことを示している。ただし、(1)さらにデータを増やして適用性を確認すること、(2)言語モデルを詳細化すること、(3)類似度計算対象を拡張すること(メッセージ間、場面や状況の取り込み)、が今後の課題である。さらに、(4)本手法の自然な話し言葉翻訳への応用を検討していきたい。

謝辞

本研究の機会を与えて下さった山崎社長、指導いただいた飯田室長、また数々の貴重なご意見をいただいた研究室の研究員諸氏に感謝いたします。

参考文献

- [1] M. Nagao, "A Framework of a Mechanical Translation between Japanese and English by Analogy Principle," in *Artificial and Human Intelligence*, eds. A. Elithorn and R. Banerji, North-Holland, pp. 173-180 (1984)
- [2] S. Sato, "Example-Based Machine Translation," Doctoral Thesis, Kyoto University (1991)
- [3] 隅田 ほか, "英語前置詞句係り先の用例主導あいまい性解消" 信学論 Vol. J77-D-II No.3 pp.557-565 (1994)
- [4] 古瀬 ほか, "経験的知識を活用する変換主導機械翻訳" 情処学論 Vol. 35 No. 3, pp. 414-425 (1994)
- [5] Y. Sobashima, O. Furuse, H. Iida, "A corpus-based local context analysis for spoken dialogues," *Speech Communication*, Vol 15, pp. 205-212 (1995)
- [6] A. S. Hornby, *Verb Patterns in "Oxford Advanced Learner's Dictionary of Current English"*, published by Kaitakusha in Japan (1974)
- [7] 花本金吾, F. J. Edamatsu, "実用英会話必携," 全面改定版, 旺文社 (1993)
- [8] 斎藤なが子, "よく使われる英検3級レベルの会話表現 333," 旺文社 (1993)
- [9] O. Furuse, et al., "Bilingual corpus for speech translation," *Proc. of AAAI -94 Workshop on "Integration of Natural Language and Speech Processing"*, pp. 84-89 (1994)