

## 音声理解システムにおける言語理解処理の シミュレーション

小田 まり子    小田 誠雄<sup>†</sup>    横田 将生<sup>‡</sup>

久留米工業大学    † 福岡工業短期大学    ‡ 福岡工業大学

我々は、内容的に妥当であれば、発話者の表現には拘らない出力を行う音声理解システムを構築している。このような処理は、もはや、音声の波動としての物理的な特徴に関する処理範囲を越えた自然言語理解処理、特に、概念処理の領域に属するものとなる。そして、不完全な音素列あるいは文字列として知覚された音声情報の理解は、自然言語理解研究における重要な課題の一つである省略語概念推定処理とほぼ等価となる。

本稿では新聞記事の文章に脱落を加えたものを音声認識部の出力と見なし、それらから言語理解部により正しい文章を推定する方法および実験結果を報告する。

## Simulation of Language Understanding Processing in the Speech Understanding System

Mariko Oda    Seio Oda<sup>†</sup>    Masao Yokota<sup>‡</sup>

Kurume Institute of Technology

† Fukuoka Junior College of Technology

‡ Fukuoka Institute of Technology

We have been constructing a speech understanding system that can infer the conceptual information which the speaker would transmit. The processing for this purpose belongs no longer to wave signal processing but to natural language understanding, especially to conceptual processing with background knowledge such as commonsense, world-specific knowledge, etc. Moreover, understanding incompletely perceived speech is nearly equal to estimating the concepts of the words omitted in texts.

For this study, an output of the speech recognition module was prepared using a newspaper article, parts of which were deliberately omitted. In this paper, we report on our method to estimate correct statements from the incomplete article in the language understanding module. We conclude with the results of our experiments.

## 1 まえがき

音声認識の理想は、話者の発話を文章として一字一句忠実に再現することである。しかし、連続音声の認識を単に音声波形の音響的な特徴や音韻規則、および構文情報のみから正確に実現することは、極めて困難である。話者が聴者に対して伝達しようとする情報は、種々のノイズが混入することによって、聴者が受けとった時には曖昧性がある概念的集合になる。従って、聴者は、この集合の中から常識や状況に関する知識に適合したものを選択することになる。このような処理は、もはや音声の波動としての物理的な特徴に関する処理範囲を越えた自然言語理解処理、特に、概念処理の領域に属するものとなる。

我々は、常識や状況に関する知識(これらを合わせて背景知識と呼ぶ)を利用した概念処理を行なうことで、音声認識の際の部分的な誤りや曖昧さを補い、音声によって伝えようとしている概念的集合が理解できるようなシステム(概念処理主導型音声理解システム)の構築を行っている [3]。

本稿では、自然言語理解処理における省略語概念推定処理を応用する事により、不完全な音声情報から尤もらしい理解結果を推定するための方法について述べる [1, 2]。また、実際に、新聞記事の文章に脱落を加えたものを音声認識部の出力と見なし、それらから正しい文章を推定する言語理解部のシミュレーション実験を行った。

本稿は、次のように構成されている。2節では、概念処理主導型の音声理解モデルについて説明を行なう。3節では、言語理解部における概念処理について説明した後、適切な理解結果を得るための方法について述べる。また、言語理解処理のシミュレーション結果とそれについての考察は4節で行う。

## 2 音声理解のモデル

我々の意図する音声理解システムは、音声認識部 (Speech Recognition) および言語理解部 (Language Understanding) よりなる。音声認

識部は、波動である音声を入力とし、候補となりうる単語列の集合(単語ラティス)を出力する。そして、言語理解部では、音声認識部の出力として得られた候補の中から意図的に妥当なものを選択する。我々のシステムは、特に、言語理解部の意味解釈課程の概念処理に重点をおき、未認識部を含む単語列や不完全な文章に対しても内容的に矛盾のない理解結果を推定し、出力する。従って、本システムは、内容的に妥当であれば、発話者の表現にはこだわらない書取結果を出力することもありうる。例えば、ある方言での会話、「「どさ。」「湯さ。」」に対して「「あなたはどこにいくのですか。」「私は風呂へいきます。」」というように、入力音声の物理的特徴を無視したような表現で理解結果を出力したりする。また、複数個の理解結果は、理解容易性に基づいて序列化する。

## 3 省略概念補完による言語理解

音声理解システムにおける言語理解処理は、音声認識部より得られる離散的な単語が必ずしも正しくないという点を除けば、自然言語理解処理における重要な課題の一つである省略語概念推定処理とほぼ等価となる。以下では、我々が提案している音声理解システムにおいて、音声認識部で不完全に認識された文章が、言語理解部に渡された場合の処理の概略を説明する。

### 3.1 未認識部分の概念推定

言語理解部では、離散的に認識された単語を全て正しいものと仮定して、その単語概念間の欠落した単語(句)概念を陽に記述することが行われる。そのためにシステムは背景知識を利用する。その場合、全ての語の意味記述が与えられているとの前提に立っており、離散的に認識された単語の意味構造がそれを含む文脈の意味構造に矛盾なく統合された時にその理解結果は成功したことになる。このことは、式(1)における推理記号( $\vdash$ )の右辺(すなわち、欠落語を補完した構造の意味解釈)が導出されるような処理を行うことを意味する。ただし、この式で  $B$  は背景知識である。表層的には式(2)が

示すように、離散的に単語認識された音声情報  $P$  の欠落箇所  $x_i$  へ挿入される語句  $p_i$  の集合 (代入)  $\theta$  を求めることになる。

$$I(P[x_1, \dots, x_n]) \wedge B \vdash I(P[p_1, \dots, p_n]) \quad (1)$$

$$I(P) \wedge B \vdash I(P\theta), \theta = \{x_1/p_1, \dots, x_n/p_n\} \quad (2)$$

このようにして得られた理解結果は、通常、複数個存在する。これらの理解結果は仮説の域を出ておらず、何らかの評価関数を用いて適切度を判定することが望ましい。式 (3)~(5) は、これらの事柄を形式化して表現したものである。

$$h(P, B) = H \quad (3)$$

$$H = \{P[x_1, \dots, x_n] \mid \theta \text{ 仮説的修復単語列}\}$$

$$= \{H_1, \dots, H_m\} \quad (4)$$

$$e(H) = H' \quad (5)$$

ただし、ここで  $h$  は仮説生成関数、 $H$  は仮説の集合、 $e$  は適切度評価関数、 $H'$  はある評価基準に基づき順序付けされた  $H$  である。

### 3.2 理解結果の適切性の評価

我々は複数個得られた理解結果、すなわち、仮説の集合  $H$  の要素に、ある観点から優先順位を付け、順序集合  $H'$  を生成する一つの方法を採用している。この方法は、 $H$  の各要素の意味理解に要するコストを計算し、そのコストの小さい順に優先順位を付けるというものである。そして、システムにとっては、そのような意味理解に要するコストの大小が、理解に関する難易度と比例するとしている。すなわち、「最も理解しやすいものが最適な理解結果である」と仮定することになる。

我々のシステムでは、表層構造 (“毒グモ、駆除、殺虫剤、散布”) のような離散的単語列より得られる意味解釈結果を述語論理式で表現している。また、表層構造から概念構造を得る際の仲介として、システムは表層依存構造 (surface dependency structure) を利用する。従って、意味理解に要するコスト (複雑度) は、表層構造から表層依存構造を経て、意味構造 (述語論理

式) に至る各処理過程におけるコストの総和として表現されるべきであろう。このような考えに基づき各過程のコストを分析してみる。

#### 1. 表層構造・表層依存構造間の変換コスト

表層構造から表層依存構造へ変換するコストは、大まかに出力文章を構成する単語の数 ( $N_o$ ) に比例すると考える。

#### 2. 表層依存構造・意味構造間の変換コスト

出力文章を構成する単語の意味記述に基づく意味構造生成処理では、主に、述語論理式間における統合処理 (ユニフィケーション) の発生回数 ( $U_o$ ) が理解に要するコストの対象になる。また、省略概念補完処理では、統合処理の発生回数 ( $U_s$ ) に加えて、補完概念を挿入する回数 ( $N_s$ ) がコストの対象となる。

以上の考察結果に基づき、意味理解に要する総コスト ( $C_T$ ) は式 (6) で与える。

$$C_T = N_o + U_o + U_s + N_s \quad (6)$$

表層的には、 $U_o$ 、 $U_s$  および  $N_s$  は、それぞれ依存構造における依存関係 (係受け対) の個数 ( $D$ )、挿入単語数 ( $W$ )、および挿入自立語数 ( $E$ ) にほぼ対応している。すなわち、 $C_T$  は、近似的に式 (7) で与えられる。

$$C_T = N_o + D + W + E \quad (7)$$

更に、依存構造の性質 ( $D = N_o - 1$ ) から、 $C_T$  は次式 (8) で近似できる。

$$C_T = 2N_o + W + E \quad (8)$$

## 4 言語理解処理シミュレーション

我々のシステムは、音声認識の結果得られるような不完全な文章 (離散的単語列) の入力に対して、概念を補完し、意味的に矛盾のない複数の結果を理解結果として得る。例えば、我々のシステムでは、表層構造 “毒グモ, X1, 駆除, X2, 殺虫剤, X3, 堺市, X4, 職員, X5” から図 1 のような表層依存構造を仲介として図 2 のような意味構造を得る。



図 1: 表層依存構造の例

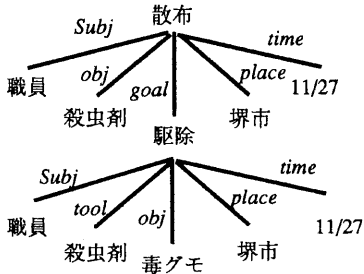


図 2: 理解結果の意味構造の例

また、システムは、これらの理解結果に対して3.2節で述べた評価関数を用いて適切度の評価を行い、優先順位の高いものから順番に出力を行う。

#### 4.1 概念補完処理

本システムは事物の概念を記述した辞書を持っており、言語理解部で理解結果を出力する際、これを背景知識として利用する。そして、入力された離散的な単語の持つ概念、および補完しようとする概念との間で矛盾が生じないかどうかを判別する。そして、入力された語の持つ概念と補完された概念が全て矛盾なく統合された場合に、理解処理に成功したとする。

物概念の場合、その物自身が持つ性質や他の事物との関係などが知識として記述されている。

例えば、「毒グモ」であれば、

上位概念：クモ(節足動物)

形状に関する知識：大きさは1mmから1cm程度、背中に赤い斑点がある。…

性質に関する知識：毒を持つ。側溝や墓石のすき間などの物影に生息する。人間を刺す場合がある。…

言語的知識：数助詞として「匹」を使う。…

…

等の知識が背景知識となる。

従って、[X1,'3mm',X2,'毒グモ']という入力に対して概念処理を行うとシステムは3mmと毒グモの大きさに関する知識から、X1=体長(が)、X2=φ(の)という語を補い、「体長3mmの毒グモ」という出力結果を矛盾なく出力する。しかし、[X1,'3m',X2,'毒グモ']という入力に対しては、「3m」と「毒グモ」の間は大きさでの関係付ができない。そのため「長さ3mの側溝に住む毒グモ」というように両者を統合するための別の概念が補われる。

また、「散布」などの事概念の記述は、概念そのものの記述以外に、事象を構成する要素との関係の記述が重要となる。例えば散布の対象は「薬剤」であり、これを行うのは「人間」でなければならない、等である。

#### 4.2 出力結果の考察

表1から表5は、新聞記事の文章に助詞などの脱落を加えたものを音声認識部の出力と見なして言語理解部に入力した場合の出力結果を示している。これらの出力結果は、[]によるネスティングで依存構造も表示している。なお、表中の\*\_\*は挿入(補完)された単語であることを意味している。また、破線より下は冗長な出力結果である。

表1は、['毒グモ',X1,'1000',X2,'匹',X3,'超す',X4] (原文：毒グモ、1000匹超す)を入力した場合の出力結果である。この場合、最適と思われる結果が最小のコスト( $C_T=14$ )で得られる。また、表2に示したように、['毒グモ',X1,'1000',X2,'人',X3,'超す',X4]を入力した場合は、「毒グモを駆除する100人を超す職員」や「毒グモを発見する100人を超す職員」などが結果として得られる。この場合の最小コストは $C_T=25$ であり、表1の例に比べて、コストが非常に大きい。これは、「毒グモ」と「1000人」の間で統合処理を行う際に、他の概念を補完する必要があったためである。また、この場合、「毒グモが100人を超す」のような数詞とそれに対応する物との関係が矛盾しているものは除去される。

表3は、['泉北郡忠岡町',X1,'同日',X2,'発見',X3] (原文：泉北郡忠岡町でも、同日発見)を

入力した場合の出力結果である。この場合の出力結果では、最適と思われる二つの結果(No1,2)が最小のコスト( $C_T=26$ )として得られている。現時点では、No.1 および No.2 のどちらかを最適結果として選択することはできない。今後、話題性などによって補充概念の表出順序も推定する必要があると考えられる。

表 4は、['毒グモ',X1,'300',X2,'匹',X3,'発見',X4,'浜寺公園',X5] (原文:毒グモが300匹発見された浜寺公園) に対する出力結果である。この場合の出力結果では、最適と思われる結果(No.1)が最小のコスト( $C_T=27$ )で得られている。また、"職員が発見する職員"のような冗長な出力は優先順位が下げられている。

表 5は、['毒グモ',X1,'駆除',X2,'殺虫剤',X3,'散布',X4,'堺市', X5,'職員',X6](原文:毒グモ駆除のため殺虫剤を散布する堺市の職員) に対する出力結果である。この場合の出力結果では、最小のコスト( $C_T=27$ )で得られている結果は、意図を表す「ため」という語が含まれていない。全ての意味構造を陽に表している最適結果は、No5 であるが、意図を表す語「ため」と勤務という概念を補完しているため、コストは大きくなる。

## 5 むすび

音声理解を目的とした概念処理の理論的モデルを提案し、計算機シミュレーションを試みた。今回達成された事柄は以下のようにまとめられる。

### 1. 概念処理モデルの提案と動作の確認

音声認識部の出力を不完全な文章とみなし、言語理解部における背景知識に基づく概念処理によりその内容を推定する方法を提案し、そのシミュレーションを行った。

### 2. 選択公準の効果の確認

言語理解部において、無意味、矛盾または冗長であると判定された解釈結果を主に排除する目的で選択公準を導入し、その効果を確認した。

### 3. 文生成過程の動作の確認

理解結果である概念構造から表層依存構造(係受け構造)を経て日本語文を生成する処理過程を実現し動作の確認をした。

### 4. 理解容易性の提案と有効性の確認

理解結果を表層構造に変換したものの(出力文)から表層依存構造を経て意味構造に至る各処理過程におけるコストの総和と逆比例の関係で理解容易性を定義した。

今後の課題としては、以下のようなものがある。

#### 1. 選択公準の充実

今回適用した選択公準は、全て、文脈内容との整合性を重視した論理的観点からの排除型のものであり、積極的に採用する型の選択公準も用意する必要がある。

#### 2. 選択公準の適用条件の定量化

選択公準を絶対的でなく、無意味度や冗長度などの程度により適用を緩和するような機構も、次の段階では必要となる。

#### 3. 音声認識部との接続

現時点において、言語理解部と音声認識部との独立をモデル化しているが、将来は、統合的なモデル化が必要となろう。この場合、入力音声の特徴(継続時間や音律など)も積極的に取り入れる。

## 参考文献

- [1] 横田将生, 小田まり子, 小田誠雄, "不完全文書処理としての音声理解", 情報処理学会自然言語処理研究会資料, 95-NL-106, pp.43-48, (1995)
- [2] Masao YOKOTA, Mariko ODA and Seio ODA, "Speech Understanding Driven by Conceptual Processing", Proc. of ICPhS 95, Vol.4, pp.324-327 (1995)
- [3] 小田まり子, 小田誠雄, 横田将生, "概念処理主導型音声理解システムの枠組について", 電機関係学会九州支部連合大会, vol.1141, (1994)
- [4] 横田将生, "心像意味論に基づく日本語接続名詞の意味解釈", 信学論, J78-D-II-11, pp.1657-1668, (1995)

表 1: [ '毒グモ',X1,'1000',X2,'匹',X3,'超す',X4] に対する出力結果

No	Result	N <sub>0</sub>	W	E	C <sub>T</sub>	P
1	[[[毒グモ],*が*],[[1000,匹,[]],*を*],超す]	6	2	0	14	1

表 2: [ '毒グモ',X1,'1000',X2,'人',X3,'超す',X4] に対する出力結果

No	Result	N <sub>0</sub>	W	E	C <sub>T</sub>	P
1	[[[毒グモ],*を*],*駆除*,*する*],[[1000,人,[]],*を*],超す,*職員*	9	5	2	25	1
2	[[[毒グモ],*を*],*発見*,*する*],[[1000,人,[]],*を*],超す,*職員*	9	5	2	25	1

表 3: [ '泉北郡忠岡町',X1,'同日',X2,'発見',X3] に対する出力結果 (一部)

No	Result	N <sub>0</sub>	W	E	C <sub>T</sub>	P
1	[[[泉北郡忠岡町],*で*],[[同日],*に*],[[*職員*],*が*],*発見*,*する*],*虫*	9	6	2	26	1
2	[[[泉北郡忠岡町],*で*],[[同日],*に*],[[*虫*],*を*],*発見*,*する*],*職員*	9	6	2	26	1
3	[[[泉北郡忠岡町],*で*],[[同日],*に*],[[*職員*],*が*],[[*虫*],*を*],*発見*,*する*]	10	7	2	29	3
13	[[[泉北郡忠岡町],*で*],[[*虫*],*を*],*発見*,*する*],[[同日],*に*],[[*職員*],*が*],*発見*,*する*],*職員*	13	10	4	40	9
14	[[[泉北郡忠岡町],*で*],[[*虫*],*を*],*発見*,*する*],[[同日],*に*],[[*虫*],*を*],*発見*,*する*],*職員*	13	10	4	40	9

表 4: [ '毒グモ',X1,'300',X2,'匹',X3,'発見',X4,'浜寺公園',X5] に対する出力結果 (一部)

No	Result	N <sub>0</sub>	W	E	C <sub>T</sub>	P
1	[[[毒グモ],*を*],[300,匹,[]],[[*職員*],*が*],*発見*,*する*],*ところの*],浜寺公園]	10	5	2	27	1
2	[[[毒グモ],*を*],[300,匹,[]],*発見*,*する*],[[浜寺公園],*で*],*散布*,*する*],*職員*	11	6	2	30	2
3	[[[毒グモ],*を*],[300,匹,[]],*発見*,*する*],[[浜寺公園],*で*],[[*虫*],*を*],*駆除*,*する*],*職員*	13	8	3	37	3
7	[[[毒グモ],*を*],*駆除*,*する*],[300,匹,[]],[[*虫*],*を*],*発見*,*する*],[[浜寺公園],*で*],*散布*,*する*],*職員*	15	10	4	44	7
8	[[[毒グモ],*を*],*発見*,*する*],[300,匹,[]],[[*虫*],*を*],*発見*,*する*],[[浜寺公園],*で*],*散布*,*する*],*職員*	15	10	4	44	7
15	[[[毒グモ],*を*],[[*職員*],*が*],*発見*,*する*],*ところの*],[[300,匹,[]],[[*職員*],*が*],[[*虫*],*を*],*発見*,*する*],*ところの*],浜寺公園]	17	12	6	52	15

表 5: [ '毒グモ',X1,'駆除',X2,'殺虫剤',X3,'散布',X4,'堺市',X5,'職員',X6] に対する出力結果 (一部)

No	Result	N <sub>0</sub>	W	E	C <sub>T</sub>	P
1	[[[毒グモ],*を*],*駆除*,*する*],[[殺虫剤],*を*],*散布*,*する*],[[堺市*役所*],*の*],職員]	11	5	0	27	1
2	[[[毒グモ],*を*],*駆除*,*する*],[[殺虫剤],*を*],*散布*,*する*],[[堺市*役所*],*で*],*勤務*,*する*],職員]	13	7	1	34	2
3	[[[毒グモ],*を*],[[*職員*],*が*],*駆除*,*する*],*ため*],[[殺虫剤],*を*],*散布*,*する*],[[堺市*役所*],*の*],職員]	14	8	2	38	3
4	[[[毒グモ],*を*],*駆除*,*する*],[[殺虫剤],*を*],*散布*,*する*],[[堺市*役所*],*で*],[[*虫*],*を*],*発見*,*する*],職員]	15	9	2	41	4
5	[[[毒グモ],*を*],[[*職員*],*が*],*駆除*,*する*],*ため*],[[殺虫剤],*を*],*散布*,*する*],[[堺市*役所*],*で*],*勤務*,*する*],職員]	16	10	3	45	5
6	[[[毒グモ],*を*],[[*職員*],*が*],*駆除*,*する*],*ため*],[[殺虫剤],*を*],*散布*,*する*],[[堺市],*で*],[[*虫*],*を*],*発見*,*する*],職員]	18	12	4	52	6
7	[[[毒グモ],*を*],*駆除*,*する*],[[殺虫剤],*を*],*散布*,*する*],[[堺市],*で*],*散布*,*する*],職員]	13	7	1	34	7
8	[[[毒グモ],*を*],*駆除*,*する*],[[殺虫剤],*を*],*散布*,*する*],[[堺市*役所*],*で*],*散布*,*する*],職員]	13	7	1	34	7
20	[[[毒グモ],*を*],[[*職員*],*が*],*駆除*,*する*],*ため*],[[殺虫剤],*を*],*散布*,*する*],[[堺市],*で*],*散布*,*する*],職員]	16	10	3	45	20
28	[[[毒グモ],*を*],[[*職員*],*が*],*駆除*,*する*],*ため*],[[殺虫剤],*を*],*散布*,*する*],[[堺市],*で*],[[*職員*],*が*],*発見*,*する*],職員]	18	12	4	52	28
35	[[[毒グモ],*を*],[[*職員*],*が*],*駆除*,*する*],*ため*],[[殺虫剤],*を*],*散布*,*する*],[[堺市*役所*],*で*],[[*虫*],*を*],*駆除*,*する*],職員]	18	12	4	52	28