

スケジューリングタスクにおける セマンティックフレームの音声認識への応用

葭原 裕久 樽松 明

電気通信大学 電子工学科 情報伝送学講座
{yoshi,kure}@apple.ee.uec.ac.jp

1 はじめに

ここ数年の飛躍的な技術の進歩に伴い演算処理装置の高速・高音声認識は認識性能が向上して単語数を限定したり話す文の組合せをある程度文法もったものに限定しておけばある程度の認識率を保てる事が知られている。しかし、不用な単語の発音や ill-formed な文構造をもった発話である自由発話の音声に対してはまだ単語認識率が十分高いとはいえず、したがって自由発話文の音声認識は、文全体の単語列を正確に認識するのは至難な状況にある。

自由発話の音声認識に対処する手段の一つとして発声の単語をすべて忠実に認識するのではなく、発声の文の意味を抽出することが考えられる。自然言語文の意味抽出の研究では、格助詞にもとづく述語や関連する語の意味抽出がある [1] が、話し言葉音声では、助詞が省略されたり発話されても弱い音のため正しく認識されない事が多いためこの方法は適さない。これまでに音声入力にケースフレームを適用したセマンティックケースフレームの方法 [2] があるが、英語の話し言葉の文を対象にしていて、日本語の自由発話の音声に対する有効性はまだはっきりしていない。本研究では、スケジューリングタスクの自由発話対話文の意味抽出を行うことを目的として、述語を中心としたキーワードの抽出を行ない、抽出されたキーワードの属性に関する性質か属性の合致する名詞を抽出していくという手法を提案する。簡単な実験により意味抽出をおこないこの手法の有効性を示す。

2 意味抽出

自動音声認識では、音声データの A/D 変換、音素認識、単語認識などの処理を経てそこから文節、文章への認識という様に次第にマクロに認識していくという工程がある。意味抽出とは、単語

の羅列から文節や文章を組み立てる、つまりその発話の意味を単語の羅列からとらえる為に必要な工程である。具体的には与えられた文章の主語や述語、それにかかる動詞や目的語を解析することにより、その文章の意味を抽出する。

3 動詞を中心としたフレームによる抽出

意味抽出を行なうにあたって、この研究の実験では動詞を中心としたフレームを用いた意味抽出を用いている。これは、ある動詞がどのような性質の名詞を取るかという情報を辞書化し、与えられた文章の「動詞」を抜き出した後にその辞書情報から名詞を抽出する方法である。辞書項目の例を表 1 に示す。

表 1: 辞書項目の一例

```
deshi | * | でし | でし | です | 判定詞 | ||| ||| |||
ai | * | あい | 空 | 空く | 動詞 | 16 行動 | ||| ga-4 人 | ga-2670 時間 |||
awa | * | あわ | あわ | あう | 動詞 | 2 状態 | ||| ni-2443 関連 |||
gozai | * | ござい | ござい | ござる | 動詞 | 2 状態 | ||| |||
```

辞書にはひとつの動詞のとる品詞の意味クラスが複数登録してあり、またその名詞の前後にどのような格助詞がくるかも関連づけられている。意味クラスにおける意味概念は日本語語彙体系 [3] にもとづいた形で使用した。本実験において、スケジューリングタスクにおける意味抽出を行なうために設定した動詞を中心としたフレームの例を図 1 に示す。

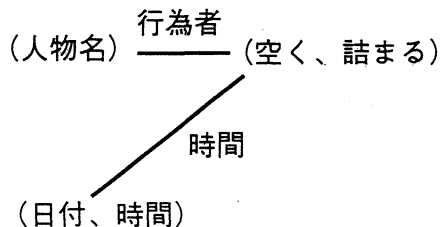


図 1: 動詞を中心としたフレームの例

4 意味関係

この研究における実験の意味関係を表 2 に示す。

表 2: 意味関係

関係	文例	表記例
of 関係	明日の三時	(三時,(of: 明日))
from 関係	三時から	(from: 三時))
to 関係	三時まで	(to: 三時))
quant 関係	明日は三時	(三時,(quant: 明日))
and 関係	一時と三時	(三時,(and: 一時))
or 関係	一時または三時	(三時,(or: 一時))

この研究の実験では、意味関係の解析を格助詞により行なっている。

上記の表 2 の「and 関係」を例にとれば、

「一時と三時」

という単語列から「と」を検出し、前後にある品詞を意味辞書と比較し、

(N: 三時,(and: N: 一時))

と抽出される。

5 実験

意味抽出に限らず、音声認識の分野ではタスクを限定してモデル化する事が多い。本論文の実験では、音声データとしてスケジュールリングタスクを扱っている。スケジュールリングタスクとは予定を決めるときに交わされる会話の事である。したがって、時間や場所などの品詞情報をいかに意味抽出できるかが、意味抽出精度に大きく影響してくる。

動詞を中心としたフレームに基づく意味抽出プログラムを用いて、抽出実験を行なった。実験条件を表 3 に示す。

表 3: 実験条件

音声データ	VERBMOBILE 音声データ
テキスト	音声データより得た 15 文
評価する文節	評価対象に当てはまる 10 文節
テストセット	10 文節中 5 文節。
オープンセット	10 文節中 5 文節
対象の動詞	頻度の高い 5 つ

スケジュールリングタスクの VERBMOBILE 会話 (No.218) から得られたデータを意味抽出プログラムのテストに用い、同データベースの他の会話 (No.201) から得られたデータを意味抽出プログラムのオープンな評価に使用した。No.218 の会話の一例を図 2 に示す。

NBN03: ええと、来週であれば火曜と金曜日一日空いた空いておりますけども、先生都合いかがですか
NBN07: ええと、もし十二日の午後でしたらばええと、一時からあ十二時から三時半まで空いておりますけれども

図 2: スケジュールリングタスクの会話 (No.218) の一例

6 意味抽出結果

実験結果を表 4 および表 5 に示す。

抽出率は次式で定義される。

$$T = \sum_{i=1}^n S_i / 3 * n$$

T : 抽出率
n : 総文節数
S : 文節

表 4: テストセット抽出率

文節名	抽出点数
NBN03A	3
NBN05A	2
NBN05B	2
NBN07A	3
NBN11A	0
合計	10
抽出率 (%)	66.7%

表 5: オープンセット抽出率

文節名	抽出点数
NBN03B	1
NBN03C	3
NBN05B	3
NBN07A	3
NBN11A	3
合計	13
抽出率 (%)	86.7%

7 考察

- 実験で使用した文章はスケジューリングタスクの中でも少数の動詞を含むものに絞ったため、局所的には高い抽出率を示した。
- ささいな認識誤りには影響されずに意味抽出が可能である。
- 日本語の動詞は変化形が多く、また品詞の基幹部分が短いものが多いため、サポートする動詞を増やすのに工夫がいる。

実験に使用した意味抽出プログラムのテスト用に用いた文節セット A の抽出率の方が、オープン評価の文節セット B より高いのは非常に興味深いところである。オープン評価に際しては、辞書に語彙を増やしただけで意味抽出プログラムに変更は加えていない。このことは辞書の追加により拡張するだけで、スケジューリングタスクにおいて有効な意味抽出を期待できるということである。

8 今後改善が望まれる点

8.1 文節の切り出し

本研究では意味抽出に用いる文節を手動で選り出した。文節の切れるキーワードにより自動的に切り出す事もまったく不可能という訳ではない。しかし、その場合はキーワードを音声認識の段階で認識できない場合や、キーワードが無い場合等もあり現実的でない。この問題の解決には、音声データのピッチの遷移やポーズや文節の時間情報から得られる韻律情報等を用いて、区分分けを行

なうことが有用であろう。この問題に関しては、他の音声認識の技術との関係、特に韻律情報などによる「文節の区切れ」や「重要な名詞に特有の韻律」が必要不可欠であると言える。

8.2 日付の結合

今回の実験では、日付のラベルにマクロ的な定義が多い。例えば、12日の場合、

jyuu ni nich

という様な3つに分かれた出力結果を音声認識プログラムから得て、それをもとに意味抽出できるのが理想であるが、真ん中の「ni」が格助詞であるのか数字の2であるのかなどの判断等、本論と関係のない所で複雑になりすぎるため、その点については人為的に

jyuuninichi

というラベルをつけて認識するように、音声認識プログラムのモデルを構築しなおしている。しかし、この様な方法を採用していたのでは「4月16日」等を抽出するためだけに膨大なラベルを辞書に用意しなければならず、音声認識の精度、速度ともに効率が悪くなる。この問題は「形態素を結合してゆき、辞書にある単語の中で、一番長い結合のものを選ぶ」という一般にワードプロセッサで使われている手法を、意味抽出プログラムで利用することにより改善できる。この場合にも、意味抽出プログラムの文法辞書は肥大化するが、それによる意味抽出の処理時間やコンピュータ資源(メモリ)の増加は大したことはなく、音声認識プログラムの方の辞書を増やして対処する場合と比べれば断然効率がよい。

参考文献

- [1] P.J. Hayes, et.al: "Parsing Spoken Language: a semantic Caseframe Approach", CMU Technical Report, (Jan. 1987)
- [2] 長尾真: 知識と推論, 岩波書店, 1988
- [3] 池原ほか: 日本語語彙体系 (岩波書店) 1997
- [4] 中筋: スケジューリングタスクにおける自由発話音声データの収集と音声認識への応用 1997