

## 音声対話システムにおける自然発話の頑健な一理解法

7R-6

肥田野勝 伊藤敏彦 山本幹雄 中川聖一

豊橋技術科学大学

## 1 はじめに

音声対話システムにおいて自然な発話における間投詞、助詞落ち、言い直し、倒置などを含む文の理解、あるいは誤認識文からの発話文の復元は対話システム品質を向上させるために必要不可欠である。本稿では人間がいかにして文の復元を行なっているかを被験者実験を通して調べ、それを参考にして復元ストラテジーを考案し、ロバストな意味理解システムを構築した。

## 2 誤認識結果の人間による復元

実際に音声認識システムで誤認識された認識結果を人間に訂正してもらった誤認識訂正実験[1]より以下のストラテジーを用いていることがわかった。

- 名詞、動詞、疑問詞などの自立語の訂正は難しい。
- 助詞の誤認識、終助詞の誤認識の修正は人間にとってそれほど難しいことではない。実際、我々は小さなタスクでは、多くの助詞落ち、助詞誤りを解析可能とする単純な手法を提案している[2]。
- 人間は名詞、動詞をできるだけそのまま使用し、文脈的、意味的にそれではおかしい場合にそれらを訂正しようとしている。
- 名詞、動詞、疑問詞の誤認識を意味理解部で修正することは、文脈を利用することによりある程度可能になる。
- 未知語、言い直し、間投詞として誤認識された重要な名詞、動詞、疑問詞をその部分の音声認識結果である不完全な音韻系列から推測し訂正している。
- 訂正困難と思われる文は重要と考えられる名詞、動詞、疑問詞から文全体を想像する。

以上のストラテジーはシステムについて全く知らない被験者による結果である。システムに熟知しているエキスパートはさらに以下の様なストラテジーも持っている。

- よく誤認識される単語のパターンを用いて訂正を行なう。また単語の音響的類似性を考慮して訂正する。

Robust natural language dialog system for spontaneous speech  
Masaru Hidano, Toshihiko Ito, Mikio Yamamoto, Seiichi Nakagawa  
Toyohashi University of Technology

単語数241、パープレキシティ174の文法で、音声認識部だけでの平均文認識率は57.4%であったが、人間による訂正によって文理解率は80%に向上する。エキスパートによる訂正の場合は文理解率は87%まで上昇する。

## 3 音声認識結果の意味理解システム

## 3.1 意味理解システムの概要

本節では、前節の人間の誤認識復元能力の観察に基づき、誤認識を含む自然発話の音声認識結果を理解するシステムを構築した。

構文解析・意味理解部は文節解析を行なった結果を係り受けに基づく文節間の依存関係を解析する。解析の途中結果はチャートデータベースに格納され、一度行った部分解析結果を保存するようになっている。また、助詞落ち、助詞誤り、倒置に対応するためにいくつかのヒューリスティックスを用いている[3]。それから係り受け解析が成功した場合は、構成性原理を使って再帰的に文節の意味表現を組み合わせて文の意味を作る。意味表現は文脈的な観点からおかしなところがないかどうかチェックされる。これをフィルタリングと呼ぶ。詳しくは3.3節で述べる。さらに、ボトムアップに意味表現を得ることができない場合は、キーワードによる意味の抽出を行う。パターンに記述された制約に適合する単語を探すことにより、全体の意味表現を得る。詳しくは3.4節で述べる。全体の処理は次のような手順で行われる。

1. 以下の処理を順次行っていき、解析が成功した時点で(2)へ行く。すべての処理で失敗した場合は(3)へ行く。
  - (a) 助詞落ち、倒置を禁止して解析
  - (b) 助詞落ちを許可して解析
  - (c) 助詞落ち、倒置を許可して解析
  - (d) 助詞の誤りを認めて、倒置を許可して解析  
(助詞が誤っていると仮定した助詞は省略されたと見なし、助詞落ちを解析するヒューリスティックスを用いる。)
2. 文脈的な知識によって、正しい内容かどうかをチェックする(フィルタリング)。
  - (a) 正しくない場合  
修正用のヒューリスティックスがある場合は、それを適用し、解析結果とする(解析終了)。修正用のヒューリスティックスがない場合は、(3)へ行く。
  - (b) 正しい場合  
得られた意味表現を解析結果とする(解析終了)。
3. 部分解析結果を用いてキーワード解析を行い、その結果を解析結果とする。

### 3.2 タスクの情報を利用したフィルタリング

フィルターは、認められない意味表現を記述したパターンとして登録されている。修正して正しく変形することが可能なものは、修正手続きをパターンと共に登録してある。このパターンに一致した場合、意味表現はリジェクトされ、修正手続きがある場合は、その手続きが適用される。パターンの例を図1に示す。patternの後にはマッチング用のパターン、modify-funの後には修正用の関数が記述される。例えば、図1のfilter1の例では、「ある」または「かかる」が主動詞の発話は質問であるはずである」という知識を記述している。

```
filter1: (pattern: ((ある かかる) (form assert))
          modify-fun: (change 'form yn-q))
```

図1 意味表現フィルターの記述例

### 3.3 キーワード抽出による意味解釈

上記までで述べた意味解釈の方法がすべて失敗した場合、トップダウン的な意味解釈を行う。各文節の意味表現を抽出し、その意味表現とマッチする変数(キーワード)を持つパターンを使って意味表現を作る。図2にこの手法の知識(キーワードパターン)の例を示す。prototypeの後には結果としての意味表現のもととなるパターン、bindingの後には変数の束縛条件である。変数の前には「?」が付いてある。図2では?aruという変数は「(ある)」という意味を持つ文節の意味表現に束縛される。その値が、パターンの変数の値として展開される。図2のキーワードパターンは「富士山にどんなホテルがありますか」などの文に対応する意味表現を抽出するもので、「富士山」などの場所の名詞、「ホテル」などの施設の名詞、「ありますか」などの「ある」という意味を持つ動詞が、音声認識結果に含まれていれば、強制的に上記のような意味表現に変換する。

```
(prototype: (?aru (form wh-q) (target (obj))
             (obj ?org)
             (at-loc ?loc))
(binding: (?aru (imi (ある)))
          (?org (sem-features org))
          (?loc (sem-features loc)))
```

図2 キーワードパターンの例

## 4 システムの拡張

今回は実験によって得られた人間の修正ストラテジーの一部しか実現しなかった。その他に以下のようなストラテジーが有効であると考えている。

- 音韻情報の利用。  
入力単語の認識部の結果が不確かなために入力を棄却しその部分を未知語として出力した場合と本当に音声認識部にとって未知語であった場合の二通りがある。我々のシステムの音声認識部の辞書と文法は、その結果に対して理解する言語理解部の辞書と文法とは異なるものを使っている。このため音声認識部では未知語である単語でも、言語理解部では未知語でないものもありうる(言語理解部の方が単語辞書を大きくしておく場合)。このため、音声理解部で未知語となったものは言語理解部で復元できる可能性がある。そのため未知語、間投詞として認識された音韻や誤認識された単語の音韻情報を利用して元の単語に復元する。これによって重要な名詞、動詞、疑問詞などを復元できる可能性がある。
- 誤りやすい誤認識パターンの利用。  
良く間違われる誤認識のパターンをシステムに与える。これによって上記の場合と同様に重要な名詞、動詞、疑問詞などの復元の可能性がある。
- 動的文脈の利用。  
今は静的な文脈のみを利用しているが、訂正実験では人間は以前になされた発話も非常に重要視している。これらの動的文脈の使用も考えている。
- N-bestの利用。  
今は認識結果は最適と考えられるもののみを対象にしているが、いくつかの認識候補結果を利用するN-bestもロバストなシステムには有用であると思われる。

## 5 むすび

自然な発話を許す対話システムを構築するために、実際に人間に音声認識部の誤りを含む認識結果を訂正してもらいその訂正ストラテジーを対話システムにインプリメントした。今後は前章で考案した方法を順次試みていく予定である。

## 参考文献

- [1] 伊藤, 大谷, 肥田野, 山本, 中川:「事前説明によるシステムへの入力発話の変化と認識結果の人間による復元」, 情報処理学会, 音声言語情報処理研究会, 94-SLP-4-7(1994.12)
- [2] 山本, 小林, 中川:「音声対話文における助詞落ち・倒置の分析と解析手法」, 情報処理学会論文誌 Vol.33, No.11, pp.1322-1330(1992).
- [3] 山本, 肥田野, 伊藤, 甲斐, 中川:「自然発話の意味理解と対話システム」, 情報処理学会研究報告, 94-SLP-2-13, pp.91-98(1994.7)