

## 情報検索対話システムにおける事例に基づく意味解釈 Example-Based Semantic Interpretation for Japanese Dialogue System

尾崎 光芳<sup>†</sup>  
Mitsuyoshi Ozaki

岡 夏樹<sup>‡</sup>  
Natsuki Oka

### 概要

今日利用可能な対話システムの技術は、語彙や文法がある程度制約されている条件でのみ実用的であり、ユーザの自由な発話を許すような音声認識処理や自然言語処理を実現することは困難である。しかしながら、ユーザにとってより使いやすいシステムを構築するためには、ユーザとシステムの会話をより人間同士に近い自然なものとすることが求められている。自然な対話の実現のために、対話システムはユーザの意図を把握して適切な対応を行う必要があり、その技術として文体や発話のタイミングなどに制限のない対話、すなわち自由対話の実現が求められると考えられる。本稿では、自由対話を実現するための要素技術のうち入力文の意味解釈を取り上げ、事例ベースの手法を用いた意味解釈を行うシステムを構築し、実験と評価を行った。

### 1. はじめに

近年、自然言語による対話システムへの社会的関心の高まりとともに、様々なシステムやその実現方法の研究が進められている。自然言語によるタスク指向の対話システムでは、ユーザの入力を解釈して、状況に応じて必要な情報を抽出する能力が求められる。しかしながら、文体や発話のタイミングといった制約のない発話を許容できる対話システムは、[1] などで研究が進められているものの、まだ実用段階ではない。本稿では、タスク指向対話である情報検索対話を対象とし、その中でユーザの発話の意味を正しく解釈できるシステムの構築を目的とする。

### 2. 事例ベースによる手法

以前より研究が行われてきた規則ベースによる対話システムが抱える問題点の一つとして、意味解釈のための規則や辞書が巨大化し、それらの改良・保守に多大な労力が必要となるという問題が挙げられる。これを克服するために、事例を用いる手法が提案されている [2]。

この手法の基本的な考え方は、発話文事例のデータベースを用意し、入力文を受け取ったときにその文と最も類似度の大きい事例を模倣することによって、規則を用いることなく意味解釈処理を行おうというものである。

規則ベースによる手法では、従来より意味解釈処理において、単語に対しては概念の階層構造を定義することにより文の意味を決定することが行われてきた。しかしこの方法には問題点が指摘 [3] されており、本稿ではこの問題を解決するために事例に基づく意味解釈の手法を採用する。

<sup>†</sup>京都工芸繊維大学大学院工芸科学研究科, Graduate School of Engineering, Kyoto Institute of Technology

<sup>‡</sup>京都工芸繊維大学, Kyoto Institute of Technology

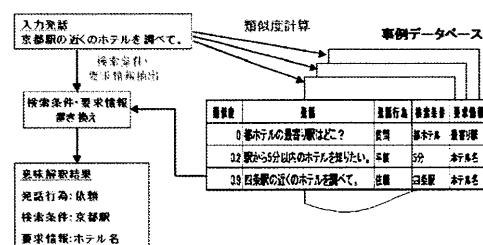


図 1: 意味解釈処理の例

事例に基づく手法では、意味解釈のために抽象化された規則を作る必要がなく、また多くの事例を用意することによって複雑な文でも正しい意味解釈を行えることが期待されるが、一方で精度の向上のために莫大な事例数が必要なことや、適切な類似度の計算方法を定めることができ難であるという短所が存在する。これらの問題の対応策として、本稿では以下の考え方を用いた。

- 事例を学習により順次拡充していくことで、段階的に性能の向上を図ることができると考えられる。
- 入力文と事例との類似度を計算する際に、文節に重み付けすることによってより精度の高い類似度を算出できるのではないか。

事例の拡充と、[1] で今後の課題とされた文節への重み付けによる意味解釈の精度の比較検討を行った。

また本稿では、基本的な部分としては従来の事例ベースの手法を採用するが、発話内におけるタスクに依存する情報を規則ベースの手法を採用して抽出を行うことでより柔軟性の高い意味解釈性能の実現を目指した。

### 3. 意味解釈手法

#### 3.1 手法の概要

本稿では、タスクを京都市内のホテル検索に設定している。用いる事例は、入力文を形態素・構文解析した構文情報を入力とし、それに対応する意味表現を出力とする組を一つの事例とする。

意味解釈の例を 図 1 に示す。以下の手順により得られた意味表現を入力発話の意味表現とする。

- 入力発話と事例内の各発話との類似度を計算する
- もっとも類似度の大きい発話に付与された意味表現を取得する
- その意味表現に対し、検索情報および要求情報を入力発話から抽出したものに置き換える

### 3.2 学習について

事例ベースでの対話システムにおいて意味解釈処理を実現するためには、

- (1) 係り受けの事例
- (2) 類似度の計算式

が必要である。このうち(2)についてはあらかじめシステムに組み込んでいるとすると、(1)を拡充していくことが意味解釈の精度を向上させることにつながる。これを実現するために本稿ではブートストラップの手法[4]を用いる。ブートストラップは事例をデータベースに順次拡充していく方法であり、以下に示す手順で拡充を行う。

1. システムに入力を与え出力を得る。
2. 出力が正しければ4.へ。
3. 出力が誤っていれば人間が正しい出力を与える。
4. 正しい入出力ペアを事例として追加する。
5. 1.へ

## 4. 実験

### 4.1 実験方法

本稿の手法を評価するために、前章で述べた意味解釈を実行するシステムを構築し、ホテル検索タスクにおける発話の意味解釈実験を行った。実験に用いる発話文として、被験者2人(20代・男性)にタスクの内容を伝え、ホテル検索のための発話を被験者がキーボードにより入力し、150発話を収集した。収集した各発話文に対応する構文情報をシステムにより形態素・係り受け解析し作成した。また、対応する意味表現の要素のうち、発話行為については筆者が人手で、検索情報と要求情報についてはシステムにより抽出し作成した。

以上により作成したデータに対し、それをランダムに10個のブロックに等分割し、1ブロックずつ訓練数を増加させながら残りのブロックをテストして意味解釈を行った。また、本実験では文節への重み付けの有無および重み付けの内容によって意味解釈の性能の比較検討を行った。なお、形態素解析には茶筌[5]を、係り受け解析には南瓜[6]を使用した。

#### [文節の重み付け]

本実験では、ある文節の重み  $w$  ( $0 \leq w \leq 1$ ) を以下のように定義した。

$$w = 1.0 - kc$$

$k$  は重み係数、 $c$  は構文木の根からその文節の節もしくは葉までの節をたどる回数である。 $k$  の値を変化させることによって、文節の重み付けによる意味解釈処理の性能の比較を行う。ただし、 $1.0 - kc$  の値が負となるときは  $w = 0$  とする。また、 $k = 0$  のとき  $w$  は定数となるので、「重み付けなし」と定める。

### 4.2 結果

前節の方法に従って行った実験結果を図2に示す。

実験を行った発話に対し、全発話のうち9割を訓練に用いたとき重み付けを行わない場合に73.3%、行った場合に80.0%の正解率を得た。このことから、文節そのも

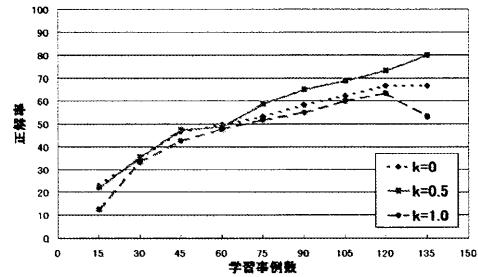


図2: 実験結果

のに重み付けを行うことが発話文の意味解釈において有効であることが分かった。

また、重み付けに  $k = \{0, 0.1, 0.25, 0.5, 1.0\}$  の5通りを用いて実験を行ったところ、 $k = 0.5$  のときに正解率が最大となった。 $k = 0.5$ 、すなわち発話文における述部と、それに直接係り受け関係を持つ文節のみを用いて類似度の計算を行った場合に意味解釈の精度が最も高くなるということが示された。この結果から、述部およびそれに直接係り受け関係を持つ文節が、意味解釈において重要であることが分かった。

### 5. おわりに

本稿では、形態素・係り受け関係を用いた事例に基づく発話文の意味解釈について、事例を順次拡充していくことによる精度の向上を確認した。また、文節への重み付けの有無による精度を比較し、本実験で用いた重み付け手法が有効であるという結果を得た。この要因として、主に語尾から質問・依頼などの文の分類がなされるため、語尾の含まれる述部に大きく重みを付けることが精度向上に繋がったと考えられる。また本稿において、規則を用いて発話文からタスクに関連する情報を抽出することを試みたが、精度向上には多くの規則の記述が必要となるため、その手法についてさらなる検討が必要である。

今後の課題として、より適切な類似度計算および重み付けの検討、応答部を実装しインタラクションを通じての対話性能評価などが挙げられる。

### 参考文献

- [1] 木村ほか：車内音声対話システムのための事例に基づく発話意図推定、情報処理学会研究報告 SLP40-20, pp.115-120, 2001
- [2] 佐藤：MBT1：実例に基づく訳語選択、人工知能学会誌, Vol.6, No.4, pp.128-136, 1991
- [3] 野口ほか：対話文脈への入力文の意味の位置付け手法、第17回人工知能学会全国大会, 2003
- [4] 長尾ほか：自然言語処理、岩波書店, pp.520-521, 1996
- [5] 松本ほか：形態素解析システム 茶筌,  
<http://chasen.aist-nara.ac.jp/>, 1997-2004
- [6] 工藤ほか：係り受け解析器 CaboCha/南瓜,  
<http://cl.aist-nara.ac.jp/~taku-ku/>, 2001-2004