

## 模擬対話データのデータベース化と事例ベース意味解析の適用

荒木 雅弘      東 郁雄      田中 吾一      堂下 修司

京都大学大学院 工学研究科 情報工学専攻

〒 606-01 京都市左京区吉田本町

本稿では、本研究室で行なっている対話データ収集、対話データのデータベース化およびデータベースの応用について述べる。最初に本研究室で行なった3種類の対話データの収集方法を説明し、そのデータから得られた知見を明らかにする。次に、書き起こしたデータに品詞・意味・発話タイプ・発話意図などの情報を付与してデータベース化する際の方針と、それらの情報を関連付けて表示するブラウザの実現について述べる。最後にそのようなデータベースの利用例として、事例ベース意味解析を提案する。

## Spoken Dialogue Database and Case-based Semantic Analysis

Masahiro ARAKI      Izuo AZUMA      Goichi TANAKA      Shuji DOSHITA

Department of Information Science, Kyoto University

Yoshida Hon-machi, Sakyo-ku, Kyoto 606-01

In this paper, we describe our method of dialogue database construction and application of this database. First, we explain our experimental collection of dialogue data. Next, we show our guideline of dialogue database construction, such that part of speech tagging, semantic tagging and utterance type / intention tagging. Finally, we propose a method of case-based semantic analysis as an example of application of this database.

### 1 はじめに

近年、音声認識および書き言葉の処理においては大規模なコーパスに基づく統計的手法が中心となっている。その音声認識と自然言語処理を統合した処理、すなわち話し言葉の処理のためにも、ある程度の規模の対話コーパスが必要である。

そのような対話データを収集するために、いくつかの条件のもとで模擬対話実験を行なった。また、そのデータを書き起こし、データベース化したのでその方法についても説明する。また、このような対話データベースの応用例として本研究室で開発中の事例ベース意味解析の手法についても

説明する。

### 2 対話データ収集

ここでは、我々が行なった3種類の対話データの収集方法と、そのデータから得られた知見について述べる。タスクはグループスケジューリングと最短経路探索問題で、前者に関しては Wizard of Oz 方式によって人間-機械の音声対話を模擬したもの、および人間-人間の音声対話を収録し、後者に関しては人間-人間の電子メール対話を収録した。

## 2.1 WOZ方式による人間-機械の模擬対話

人間-機械の対話システムを作成するための基礎データを得るために、Wizard of Oz方式によって、音声対話データを収集した。Wizard of Oz方式による対話収集では、本来のシステムの機能のどこまでを人間がカバーするかが問題となるが、今回の収録では応答生成を素早く行なって、より自然な対話を実現することを目標として、システムの機能のうちの認識部から応答生成までを全て人間が行った。

収録状況を図1に示す。システム役とユーザー役は別の部屋にいる。システム役がメニューによる応答文生成ツールを用いて発話文を生成すると、そのテキストはユーザー役の前にあるワークステーションに渡され、音声合成器によって音声データに変換されて、ユーザー役のヘッドホンから出力される。ユーザー役の発話は、マイクからアンプを通してシステム役のスピーカーへ出力される。なお、被験者には実際には人間が発話を聞いて応答を生成している事は全く告げず、そのことに気付いた被験者も皆無であった。

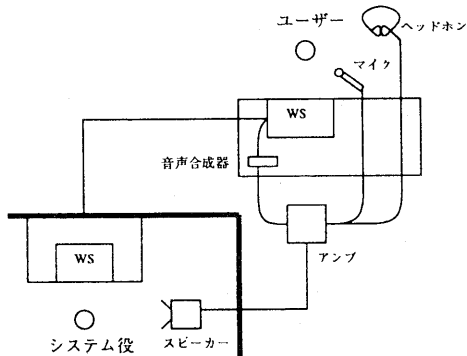


図1: WOZ方式での対話収録環境

タスクは、グループスケジューリングで、グループ会議の設定・競合する複数の予定の登録・優先度を考慮した会議室選択の3つの状況を設定し、対話を行なった。

収録したデータは5人分、25対話で、システムの発話によって中断した無効なユーザーの発話を除いた発話数は1015発話であった。システム側の発話は655発話、ユーザー側は360発話であった。図2にWOZ方式での対話例を示す。

なお、このデータに関しては発話タイプ、発話意図の出現頻度などに関する統計的データが得られている[1]。

模擬対話(書き起こし)  
 タスク: スケジュール管理  
 トピック: 会議場選択/人間-機械  
 対話番号: kyo0311  
 収録年月日: 1994/12/15  
 S/U: kyo0207(woz)/kyo0201  
 テキスト作成日: 1995/2/11

S: 堂下研対話システムです。個人スケジュール管理やグループスケジューリングを行ないます。  
 U: 御用件を、どうぞ。  
 U: 会議室のうちで、プレゼンテーションシステムが利用できる、部屋はどこか、教えてください。  
 S: プレゼンテーションシステムがあるのは、中会議室、セミナー室です。  
 U: (OHPかビデオ)、OHPとビデオを両方利用できる部屋はどこか教えてください。  
 S: 大会議室です。

図2: WOZ方式での対話例

## 2.2 人間-人間の音声対話

前節で述べたWOZ方式による人間-機械の模擬対話と比較して、人間-機械の対話に特徴的な点を抽出する目的で、前節と同じタスクで人間-人間の対話を収録した。対話の収録状況を、図3に示す。それぞれのスピーカーからは、相手の音声しか聞こえていないようにした。図4に人間-人間の対話例を示す。

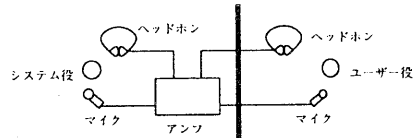


図3: 人間-人間の音声対話の収録環境

現在書き起こしが終わっていないので、人間-機械の模擬対話と比較した統計的な情報は得られていないが、データを観察すると表1に示すような知見が得られた[1]。

発話前の空白の時間に関しては、人間相手の発話ではあまり長くないのに対して、機械相手の発話ではかなり長くなる。これは、人間同士の会話ではどちらも話さない状態を少なくするためとりあえず話し始めるが、機械を相手にする場合は、ある程度発話内容を頭の中で組み上げてから発話しているためと考えられる。すなわち、音声対話システムにおいては、ユーザ発話を促す発話をシステムがするタイミングは、通常の人間-人間の対話に比べて長

模擬対話（書き起こし）  
 タスク：スケジュール管理  
 トピック：複数の予定の登録／人間－人間  
 対話番号：kyo0221  
 収録年月日：1995/1/31  
 S/U：kyo0206/kyo0203  
 テキスト作成日：1995/2/11

S: [あ]、それでは、予定の登録をお願いします。  
 U: はい。  
 S:  
 U: [えーっと]、教室会議の、  
 S: [はい]、  
 U: [はい] 日時と場所を、設定したいんですけども。  
 U: [えーっと]、  
 S: [はい]、  
 U: それで、他の教授の…予定を、知らせて下さい。  
 S: [えーと]、  
 U:  
 S: 教授全員で、  
 U:  
 S: 予定、空いている時間は、[えーっと]、月曜日の、14時から17時  
 まで、[と]、火曜日の、10時半から12時まで、  
 U: [はい]、  
 S: 同じ火曜日の、[えー]、14時半から、17時半までです。

図 4: 人間－人間の対話例

表 1: 人間－機械の対話と人間－人間の対話の違い

	人間－機械	人間－人間
発話前の空白	長い	短い
相槌	なし	頻繁にあり
確認	少ない	多い
割り込み	少ない	多い

くるとるべきであるといえる。

次に、相槌、確認、割り込みであるが、これらは人間－人間の対話には多く見られたが、人間対機械の対話ではあまり見られなかった。これは、人間対機械では発話のターンが文単位であるのに対して人間－人間ではより小さな単位であるためと考えられる。

### 2.3 人間－人間の電子メール対話

総合的な音声対話システムを構築するためには、要素技術に関する研究も必要である。我々は対話処理技術の高度化を目指して、書き言葉による対話データも収集した。収録方法は、一般にキーボード対話と呼ばれているものとは異なり、電子メールによる方法を用いた。これは以前の対話履歴全てを引用記号なしで引用し、メールの最終行に話者を表す記号と自分の発話を書いて、相手に送るものであ

る。電子メールによる方法を用いた理由としては、(1) ターンが明確なこと、(2) 考える時間が十分あるので発話が洗練されること、(3) それによってユーザプランなどの対応がとりやすいことがあげられる。

ただし、タスクはこれらの特徴が明確に出るようなものに変更した。ここでは、最短経路探索問題をタスクとした。具体的には、各被験者に(駅の連結関係や駅名の有無などに)若干の相違がある地下鉄の路線図のようなものを渡し、両者に共通の情報を用いて、出発地から目的地までの最短経路を探索するというものである。なお、このタスクは[2]で、対話システム同士の対戦によってシステムの評価を行なうために提案されたものである。図5に人間－人間の電子メール対話の例を示す。なお、カタカナ1文字で表されているものは駅名である。

電子メール対話タスク：最短経路探索  
 トピック：出発地から目的地までの最短経路／人間－人間  
 対話番号：  
 収録年月日：1995/6/21  
 A/B：/  
 テキスト作成日：1995/9/27

A: アから上へでまでいけますか。  
 B: マから左にいき、ロカルで右にいけばいけます。  
 A: ツからオを過って上へ、へまでいけますか。  
 B: いけます。  
 A: へからルへいけますか。  
 B: いけます。  
 A: オの上の駅と、レの上の駅の名前を教えてください。  
 B: オの上は、左がヌ、右がミです。レの上はヨです。  
 A: 経路は、アナマツオミエヘルニテレヨモユタイでよろしいですか。  
 B: アの上がナで、モの上がユならよろしいです。  
 A: そのとおりです。対話を終了します。

図 5: 電子メール対話の例

この方法によって、被験者6名、異なる4種類の問題に対して12対話を収録した。このデータに関しては対話の分析の他に、事例ベース意味解析用の事例または評価用データとして用いている。事例ベース意味解析については4章で述べる。

### 3 対話データのデータベース化

ここでは、対話データそのものや付加された情報をデータベース化する方法について述べる。まず、我々がどのような情報を付加してデータベース化しようとしているかについて述べ、次に元データとそれらの間の対応を視覚的に捉えられるブラウザについて述べる。

### 3.1 対話データに付加する情報

対話データに付加する情報としては、図6に示すように、品詞情報・意味情報・発話タイプ・発話意図を考えている。

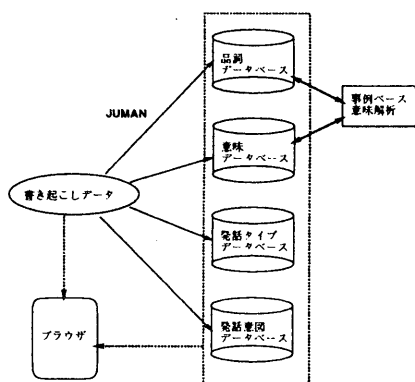


図6: 対話データのデータベース化とその応用

品詞情報に関しては、ほぼタスク独立なものが想定可能である。図2に示したWOZ方式での対話例に対するタグ情報を図7に示す。

```
<text>
<body>
<div type=section>
<U Id= Who=System>
<w id= type=znpn> 壁下 </w>
<w id= type=jb> 研 </w>
<w id= type=zn> 対話 </w>
<w id= type=zn> システム </w>
<w id= type=ha3 lemma='です'> です </w>
<w id= type=dtk> 。 </w>
<w id= type=zn> 個人 </w>
<w id= type=zn> スケジュール </w>
<w id= type=zn> 管理 </w>
<w id= type=hp> や </w>
<w id= type=zn> グループ </w>
<w id= type=zn> スケジュールリング </w>
<w id= type=hp> を </w>
<w id= type=zv2 lemma='行なう'> 行ない </w>
<w id= type=ha3 lemma='ます'> ます </w>
<w id= type=dtk> 。 </w>
</u>
```

図7: タグ情報の例

意味データベースに関しては、現在作成中であるが、ARPAのSemEvalプロジェクト[3]で採用されている述語-引数構造(Predicate argument structure)を採用する予定である。

また、発話タイプ・発話意図に関しては、永田らによって目的指向対話に現れる発話タイプ・発話意図の分類がなされており[4]、我々も基本的にその分類に従った発話のラベル付けを行なって、その出現頻度やエントロピーを求めた[1]。

### 3.2 対話データのブラウザ

これまで述べたような音声データ・書き起こしテキスト・種々のタグデータなどは、通常ディスク上の異なったファイルで保存されている。これらの情報の作成やメンテナンスには、これらを関連付けてアクセスできるツールが不可欠である。我々は、書き起こしテキストからHTMLテキストへの変換ツールおよびディレクトリ構造に基づいた参照環境作成ツールを開発し、Mosaicを用いた対話データのブラウジング環境を構築した。図8に表示例を示す。



図8: Mosaic 上でのブラウザ

## 4 事例ベース意味解析

ここでは、3章で述べたデータベースの応用例として、事例ベース意味解析について述べる。事例ベース意味解析の目的は、頑健な文解析手法の構築、すなわち非文法的な文

に対しても、何らかの解析結果を出力するというものである。入力文に対して通常の構文解析を行ない、失敗した場合に対話データベース中から類似の表現を検索して、その事例を参考に入力文の解析を行なう方法である。

事例ベース推論を用いた関連研究としては Cardie の語彙知識の獲得がある [5]。Cardie は文中の単語をタスク依存の性質を持つ open class word(名詞、動詞など)と、タスク独立の機能語である closed class word に分類し、各 open class word に対応する事例を設定した。各事例は、その open class word に関する情報(単語そのもの、品詞、一般的な意味など)、前語各 2 語に関する情報および文全体に渡る情報(主語・動詞などの一般的な意味など)から構成されている。Cardie の方法は、新しい入力文の open class word に対して近い事例を検索し、その語に関する情報を自動獲得するものである。

我々はこの方法を非文法的な文に対する意味解析手法として用いる。図 9 に事例ベース意味解析システムの概要を示す。

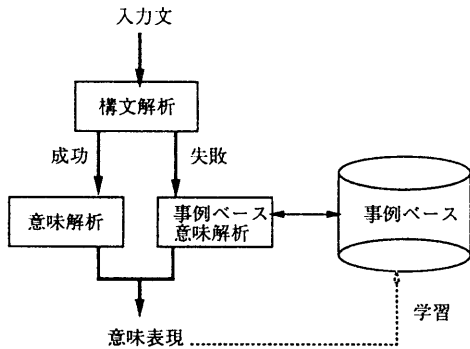


図 9: 事例ベース意味解析システムの概要

ここでの非文法的な文とはシステムが想定している文法から外れたものを意味する。今回は、タスク規模の小さい最短経路探索タスクに関して事例ベースを作成し、検索実験を行なった。現在は、意味データベースの仕様が確定していないので、表層の単語の情報と品詞の情報のみで事例を作成した。名詞、動詞、形容詞を open class word とし、1つの open class word とそれを含む文を一つの事例とした(図 10)。1つの事例は 10 個の要素を持ち、次の 2 つの特徴に分けられる。

- word definition (2) … その語自身の単語、品詞
- local context (8) … その語の前後二つずつの単語、品詞

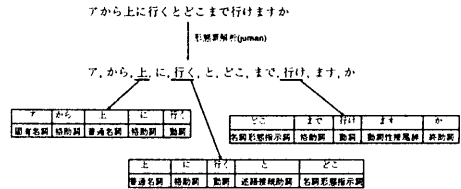


図 10: 事例

#### 4.1 事例の作成

まず、対話データベースを事例ベース(88文)と実験用に分け、事例ベース用の文を対象に DCG を記述する。事例ベース用の文は構文解析され、単語情報から事例を作成する。事例の作成過程は次のとおりである。

1. 形態素解析システム JUMAN[6]により、形態素に分割されたそれぞれの語は、品詞についての情報も持つ。文中に open class word があれば、その語に対する事例を作成する。
2. word definition の要素が自動的に埋められる。また、その語の意味情報をユーザが入力する。(今回、検索では意味情報は利用していない。)
3. 文中の全ての open class word についての word definition が定義されると、各 open class word の前後 2 語ずつの要素(local context)が埋められる。該当する語や、意味情報が無い場合 nil としておく。
4. 1つの open class word に対し、word definition, local context 及び文自体を事例として登録する。

#### 4.2 事例の検索

事例の検索法を以下に示す。

1. 入力文に対し、DCGにより通常の構文解析を行なう。構文解析が成功した場合、事例ベースの検索は行なわない。

2. 構文解析が失敗した場合、事例ベースを検索する。事例の作成時と同様に、入力文の各 open class word に対し word definition, local context の計 10 個の要素を定義し、事例ベースの要素とマッチングを行なう。
3. 要素が一致した回数をその事例の得点とし、最高得点の事例を入力文の各 open class word に対する検索事例とする。
4. 各 open class word に対する事例のうち最も得点の高いものを、その文に対する検索事例とする。

図 11 に検索例を示す。

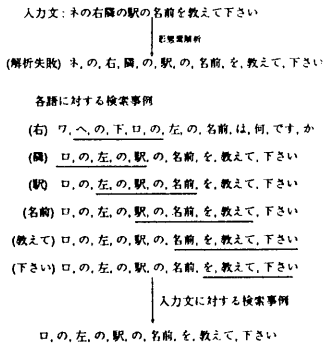


図 11: 検索例

#### 4.3 検索実験結果

88 文を対象に DCG を記述し、事例を作成した。生成された事例は 178 個であった。検索結果については、構文解析が失敗した 50 文に対し、

- 検索成功と思われるもの … 31 例
- 検索失敗と思われるもの … 19 例

であった。

この評価については、ほぼ同じタイプの文が検索されたもの、すなわちその事例を用いて意味解析を行なって成功しそうなものを検索成功とした。しかし、実際に意味解析を行なったわけではないので、この評価は正確なものではない。一般に、未知語を含む文に対して弱いようである。

## 5 おわりに

本稿では、対話データ収集、対話データのデータベース化および事例ベース意味解析について述べた。今後の課題としては、意味データベースの仕様検討及び実現、事例ベース意味解析システムのインプリメントおよびスケジューリングタスクでの評価が挙げられる。

## 謝辞

対話データのハイパーテキスト化ツールを作成して下さい京都大学工学部情報工学教室の坂田一拓氏に感謝します。また、データ収集・データベース化・分析などをお手伝いいただいた京都大学工学部堂下研究室の皆様感謝します。

## 参考文献

- [1] 東郁雄, 荒木雅弘, 堂下修司. 音声対話データの収録と意味的情報の統計的分析. 1995 年度人工知能学会全国大会講演論文集, pp. 541-544, 1995.
- [2] 橋田浩一, 伝康晴, 長尾確, 柏岡秀紀, 酒井島津 明. 対話リーグ戦: 対話システム性能評価コンテストの提案. 言語処理学会第一回年次大会 発表論文集, pp. 309-312, 1995.
- [3] Moore R. C. Semantic evaluation for spoken-language systems. In *Proc. of ARPA Human Language Technology Workshop*, pp. 126-131, 1994.
- [4] 永田昌明, 鈴木雅実. 日英対話コーパスへの発話行為タイプ付与の試みとその統計的対話モデルへの利用. 人工知能学会研究会資料, SIG-SLUD-9302-7, pp. 49-56, 1993.
- [5] C. Cardie. A case-based approach to knowledge acquisition for domain-specific sentence analysis. In *Proc. of AAAI*, pp. 798-803, 1993.
- [6] 松本裕治, 黒橋禎夫, 宇津呂武仁, 妙木裕, 長尾真. 日本語形態素解析システム juman 使用説明書 version 2.0, 1994.