

# 音声対話システムにおける 移植性の高い汎用的意味理解部の構築

梅田 将満<sup>†</sup> 小暮 悟<sup>††</sup> 中川 聖一<sup>†</sup>

<sup>†</sup>豊橋技術科学大学 情報工学系  
〒441-8580 豊橋市天伯町雲雀ヶ丘 1-1  
{umeda, nakagawa}@slp.ics.tut.ac.jp

<sup>††</sup>愛知教育大学 教育学部  
〒448-8542 刈谷市井ヶ谷町広沢 1  
skogure@auecc.aichi-edu.ac.jp

近年、音声認識技術や言語処理技術の要素技術に関しては確立しつつあり、実用化に向けシステム開発が進んでいる。実用化などを考慮した場合、今までのような使い易さ、頑健性などに関する技術だけでは不十分であり、拡張性や移植性なども十分考慮する必要がある。そこで我々は一般的なデータベースと対話例を用意するだけで音声によるデータベース検索ができる汎用性の高いデータベース検索システムを目指し、開発を進めており、今回意味理解部を改良した。以前のシステムでは、意味辞書がシステムに依存してしまい、移植性のコストが高くなってしまいう原因になっていた。そこで、システムに特化しない移植性の高い日本語意味解析を目指すために汎用的な日本語の意味体系を持ったEDR電子化辞書を利用して意味理解部を構築し、本研究室で開発されている汎用性の高いデータベース検索音声対話システムの枠組みに組み込んだ。このシステムを用いて実際に3つの検索タスクの音声対話システムを構築し、意味解析性能および移植性を評価した。

## Construction of Highly Portable General Interpreter for the Spoken Dialogue System

Masamitsu UMEMA<sup>†</sup>, Satoru KOGURE<sup>††</sup>, and Seiichi NAKAGAWA<sup>†</sup>

<sup>†</sup>Department of Information and Computer Sciences  
Toyoashi University of Technology,  
1-1 Hibarigaoka, Tenpaku, Toyoashi 441-8580, Japan  
{umeda, nakagawa}@slp.ics.tut.ac.jp

<sup>††</sup>Faculty of Education,  
Aichi University of Education,  
1 Hirotsawa, Igaya, Kariya 448-8542, Japan  
skogure@auecc.aichi-edu.ac.jp

Recently the technology for speech recognition and language processing for spoken dialogue systems has been improved, and speech recognition systems and dialogue systems have been developed to be practical use. In order to become more practical, not only those fundamental techniques but also the techniques of portability and expansibility should be developed. We already presented the portability of spoken dialogue systems. We focused on the portability of spoken dialogue system, especially, the interpreter and retrieval modules. When an application developer wants to change the domain or task, he/she modifies or prepares only domain/task dependent parts, such as semantic features, proper nouns, etc. We developed a general semantic analyzer independent of task or domain using the EDR electronic dictionary, which represents the relationship of Japanese vocabulary. We adapted the system from Mt.Fuji sightseeing guidance to a literature retrieval system and a hotel retrieval system and showed its ability of understanding and high portability.

### 1 はじめに

一般に、音声対話システムは、対話の対象となる領域・分野を限定して開発されるのが普通である。また、利用者が解決・達成したい問題・処理の種類によっても音声対話システムの仕様が異なることが多い。ここで、本稿では対話の対象となる領域・分

野のことをドメイン、「検索」「予約」のような利用者が解決・達成したい問題・処理のことをタスクと呼ぶことにする。

通常、音声対話システムの仕様は、ドメインやタスクに大きく依存するが、音声対話システム全体のモジュール性を高め、システムの移植性や拡張性を上げるためには、システムのどの部分がドメインや

タスクに依存して、どの部分がドメインやタスクからは独立であるかを明確に区別する必要がある。しかし、現実世界には様々な対話形式が存在し、そのすべてに対処するようなシステムを構築することには困難が伴う。

一般的に音声対話システムの移植性とは、システムのドメイン・タスクを変更することの容易さを示し、拡張性とは、ドメイン・タスク知識をあとから動的に追加・修正・削除することの容易さを示す。

実際に新しい音声対話システムを最初から構築するには莫大なコストがかかることから、今後、これまで開発してきたシステムを他のドメイン用に変更したり、汎用的なシステムを開発することが重要になってきている。実際、「移植性」や「拡張性」を重要視する研究も行なわれてきている。Carnegie Mellon 大学の Matthias [1] は、音声対話システムのアーキテクチャーを作る枠組みを示し、ドメイン・言語特有の知識情報を対話アルゴリズムから切り離すことにより、移植性を高めている。実験では6つのドメイン・言語固有の知識ソースでそれぞれ8~12時間以内に対話システムを構築できることを示している。

我々は、音声対話システムにおける移植性に関する研究を行なってきている。既存のシステムである富士山観光案内 [2] を東三河観光案内に実際に適用し、このような似たドメインへの変更でも30日・人かかり、全く違うドメインへの変更にはさらに大幅な作業時間がかかることを示した [3]。これらのシステムでは、音声対話システムの構成要素としては音声認識、言語理解、応答生成と対話管理が考えられるが、それぞれのモジュールについて移植性が検討されている [4]。ここでは、データベース検索タスクにおいてシステムをドメインに独立な部分(データ)、ドメインに独立な部分(データ)等に明確に分離し、例えば異なるドメインへの移植は、ドメインに依存する部分のみを変更するだけでよい設計としている。

これらの研究で問題になるのが、対話システムの一の核となる意味理解部の移植性である。意味理解部では自然言語を扱うという観点から意味辞書などドメイン・タスクに依存する部分が多く、移植の際のコストが大きくなってしまおうという問題がある。そこでドメイン・タスクに依存しないさまざまな日本語文に対応できるような意味理解部の構築を目指すこととした。すでに原田らはEDR電子化辞書に記載された情報を元に、日本語文を意味解析し格フレーム群に自動変換するシステム SAGE(Semantic frame Automatic GEnerator) [5] を開発し、意味解析システムとして実利用を開始できるレベルのものできつつある。

我々もシステムに特化しない、すなわちドメイン・タスクに独立な移植性の高い日本語意味解析を目指すために汎用的な日本語の意味体系を持ったEDR電子化辞書を利用して意味理解部を構築する手法を提案している [6]。この手法を発展させ、我々が開発している汎用性の高いデータベース検索音声対話システム [4] の枠組みに組み込む。すなわち、データベース検索タスクに対応するためのタスク依存の

データを適用し、さらに富士山観光案内ドメインのデータを用意することで、富士山観光案内ドメインにおいてデータベース検索タスクを実現し、その意味理解性能を評価する。さらに別の検索タスクとして、文献データベースとホテルデータベースの2つのドメインに移植し、その移植性および意味理解性能を評価する。

## 2 EDR 電子化辞書を用いた意味解析

従来の我々のシステムではシステムのドメイン・タスクが変わることに専用の意味辞書を用意しなければならない点が移植の際のコストが大きくなる原因となっていた。特に、動詞意味辞書(動詞の格フレーム情報)を1つ1つシステム開発者が動詞ごとに用意する手間が大きい。

そこで、ドメイン・タスクに依存しない汎用的な意味理解部を実現するために汎用的な意味辞書であるEDR電子化辞書 [7] を用いて意味解析を行うことにした。EDR電子化辞書では、階層構造的な意味体系が記載されており、これを用いた意味理解を試みる。そこで、原田らが開発している日本語意味解析システム SAGE(Semantic frame Automatic GEnerator) [5] が行っている考え方を利用し、我々のシステムで用いている動詞の格フレームによる意味解析を中心として意味理解部を構築することとする。

意味解析の前処理として形態素解析器「茶釜 2.3.3」 [8]、および係り受け構文解析器「南瓜 0.43」 [9] を利用した。これにより、各単語、各文節間の係り受け関係が得られる。これらの文節における中心語の間の関係(深層格)を決定する処理を本研究での意味理解とする。

### 2.1 概念識別子の付与

意味解析をEDR電子化辞書で行っていくの際に際して、各単語にEDR電子化辞書の概念識別子を付与する。概念識別子とは概念辞書の基本構成要素である概念に関する情報であり、6桁の16進数の数として表現され、概念の同一性を保持する目的のためのものである。この情報がEDR電子化辞書の単語辞書に各単語の意味ごとに品詞などの文法情報とともに、意味情報として記載されている。

この時、茶釜によって解析された品詞と単語辞書による品詞で一致した概念識別子のみを付与するように意味的な制約を行う。また、EDR電子化辞書では概念が詳細になるほど下層になるように木構造で体系化されている。そこで、概念体系の中でこの一致した概念識別子から木構造の根までたどったパスをその単語の概念体系として保持しておく。これは、各関係を表す辞書が中間概念で記述されているからである。以降、これら概念識別子の概念体系をもとに意味解析辞書による解析を行っていく。

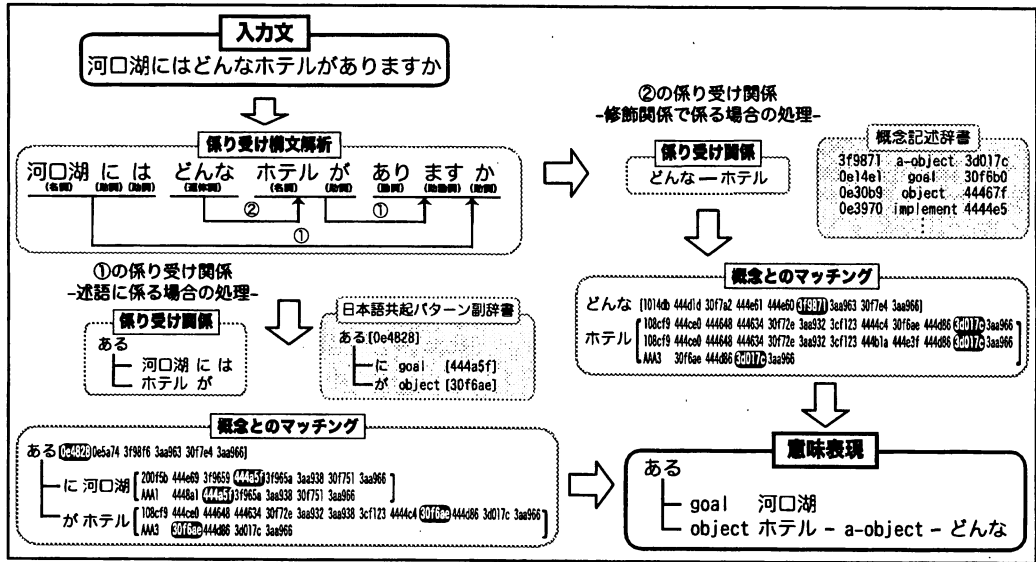


図 1: 意味解析手順例

## 2.2 意味解析辞書による解析

意味解析手順例を図 1 に示す。この例では、係り受けのパターンによって 2 種類の解析方法がある。1 目目のパターンは、従来研究の意味解析手法に基づき、動詞の格フレームの考え方をを用いた。EDR 電子化辞書には、共起辞書に付与している日本語共起パターン副辞書がある。これは各動詞（概念識別子ごと）に対して、どのような表層格（助詞）でどのような概念識別子をとって、それがどのような深層格になるのかを表した辞書である。図 1 の場合、「ある」という動詞についての格フレームの条件にしたがって、表層格のマッチングと概念識別子のマッチングが行われ、深層格が決定される。

2 目目のパターンは形容詞、連体詞などが名詞などに係るような場合である。これには EDR 電子化辞書の概念記述辞書を用いた。概念記述辞書というのは係り受け関係にある単語の 2 概念間のみによって関係が決まる辞書である。この辞書を用いて 2 単語間の関係を得ることとした。2 単語間の概念識別子とのマッチングを行い、関係のある 2 概念を見つける。図 1 では、「どんな」と「ホテル」の関係が概念記述辞書の“a-object”の行の概念識別子とマッチしたため、深層格が決定されている。

## 2.3 概念距離率による格決定

それぞれの単語の係り受け関係より、意味表現が得られる。しかし、概念識別子はその単語の意味ごとに複数存在し、また概念体系のパスも複雑に分岐することがある<sup>1</sup>。

<sup>1</sup> EDR 概念体系辞書では多重継承を許し、一つ概念が二つ以上の上位概念を持つ場合もあるので木構造にはならないが、本稿では木構造の用語を援用することもある。

また、多数ある辞書とのマッチングにより、多くの深層格のパターン（意味表現）が生成されてしまう。そこで、これらの意味表現の優先順位をつけて制限する必要がある。

SAGE[5]では深層格のとりやすさを示すのに、概念体系での距離（木構造でどれだけ離れているか）を計算し、この値の逆数をとるべき深層格ごとに足し合わせて尺度として用いる。また、Yuhua らの単語間の類似度を求める研究 [10]では、WordNet を使って類似度を定義するのに単語間の概念距離を線形的に数えるのではなく、深さによって指数関数的に類似度が増える関数によって定義している。

本稿では、概念距離率という考えを用いることとした。それぞれの単語における付与されている概念識別子から、概念構造木の根までのパスのどの部分（中間概念）でマッチしたかによる割合で表したものである。ここでは、2 単語間の各関係ごとに値が決定する。係り元の単語のある概念識別子による概念体系のパスの長さを  $L_f$ 、葉から何個目の概念識別子でマッチしたかを  $M_f$ 、同様に係り先についても  $L_t, M_t$  を定義する。これによって、2 概念間のマッチした概念識別子についての概念距離率  $P_{gai}$  が定義できる。

$$P_{gai} = \frac{M_f}{L_f} + \frac{M_t}{L_t}$$

この値  $P_{gai}$  がより小さければ、より詳細な概念識別子でマッチングしていることになる。そこで、この値が小さければ小さいほど、より詳細な意味概念を捕らえているとみなした。したがって、この 2 概念間の概念距離率  $P_{gai}$  が最小になる関係を採用することによって、関係を一意に決定できる。図 2 の場合、「ある」と「河口湖」という単語間には“goal”と“agent”という 2 つの関係が成立している。しか

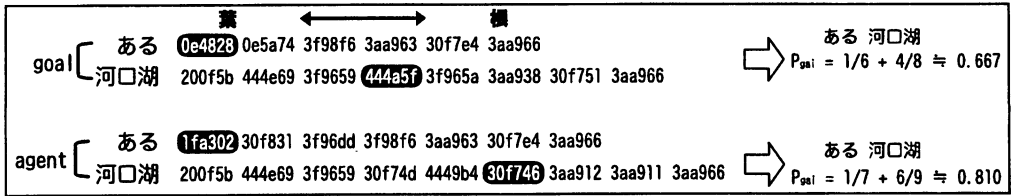


図 2: 概念距離率の求め方

しこの場合は、より概念距離率が小さい“goal”格が採用されることになる。

## 2.4 音声入力に対応した意味理解

音声対話システムを扱う場合、音声対話特有の現象に対する頑健な意味理解が必要となる。我々のシステムでは、以下のような処理を行っている。

- 助詞落ちの処理  
日本語共起パターン副辞書の格フレームパターンの意味解析手法を拡張して行う。従来法であると助詞が省略された場合は、マッチしないので格フレームが形成できない。例えば、「河口湖にホテルありますか」などのように助詞「が」が抜け落ちている場合、ホテルを“object”格として決定することができない。そこで、助詞の条件を無視して、概念識別子がマッチしていることより、助詞「が」があると仮定して処理を行う。ただし、概念距離率にペナルティとしてコストを加えて、助詞が正しくマッチする場合より優先順位を下げるようにする。今回、明らかに優先順位が下がるように明示的に+20の値を加えている（正しくマッチした場合は、約1程度の値である）。
- 訂正発話の処理  
形態素解析後に言い直しの単語とマッチングさせて、訂正している単語を棄却する処理を行った。形態素解析後の各形態素に対して、あらかじめ用意してある訂正発話キーワード列（“いや”、“じゃなくて”、など）と単語、品詞、原形の情報にマッチングさせて、そのマッチした一単語前の単語とともに棄却する処理である。
- 音声切り出し誤りによる挿入誤り除去  
音声取り込みの際、ノイズなどにより文頭・文末の音声の切り出し誤りが起きることが多くなり、その部分の認識の際、単語挿入誤りになる場合がある。そこで、形態素解析後、係り受け構文解析する前に chart-parser[11]に通すことで決められた文法での発話に制限することにより、文の前後に現れた誤認識や文法外の表現を削除することができ、文の核となる部分だけを抽出することができる。

## 3 検索タスクへの適用

今回開発した汎用的な意味理解部をある限られたドメイン・タスクへ適用することを考える。ここでは、データベースに登録されているデータを意味表現を用いて的確に検索を行うというデータベース検索タスクにおける処理を示す。

### 3.1 辞書のタスク適応

データベース検索タスクを達成するためには、検索したデータベースに登録されている単語カテゴリ（検索項目）を概念識別子として概念体系に組み込む必要がある。そこで、データベースに登録する際に用意されるデータの一覧より、カテゴリごとに得られた単語リストを利用して、カテゴリごとに概念識別子を付与する。図3の“AAA\*”のようにシステム開発者が定義した概念識別子によって各データベースのカテゴリごとにまとめられ、それをEDR電子化辞書の概念体系に組み込む。

また、この概念体系への組み込みと同様に茶釜の辞書にもこれらの単語がないと形態素解析する際に形態素（単語）として解析されないため、意味理解の際に単語が利用できなくなってしまう。したがって、茶釜の辞書にもこれらの単語を登録する必要がある。

以上の処理を行うことによって、データベース検索タスクにおける意味理解を行うための準備が完了することになる。すべての単語についてこれらの処理を手作業で行っているのは移植性の障害となってしまうが、これらはツールを用いて自動的にEDR電子化辞書、茶釜辞書へ登録することが可能である（後述する）。

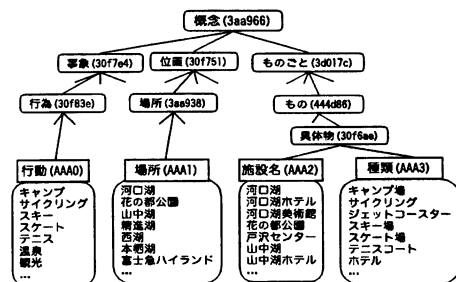


図 3: 概念体系への組み込み例

### 3.2 検索キーワード生成部

意味表現からデータベースにて検索するためのキーワードを生成する。意味表現を生成したことによって係り受け関係とその間の深層格が決定できた。これらをもとに条件を与えることによって検索するためのキーワードを生成することができる。

キーワードを生成するためにキーワード変換則というタスク依存のルールを用意する。基本的には深層格、概念識別子の条件により、どのカテゴリの条件で検索をすればよいかを定義する。

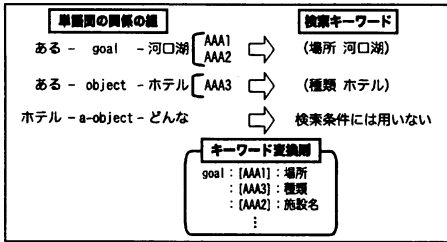


図 4: 検索キーワード生成手順の例

図 4 に、図 1 で解析した意味解析結果から検索キーワードを生成する手順を示す。各 2 単語間の関係の組を抽出して、それぞれの単語に付与しているキーワードカテゴリごとに定義した概念識別子をもとに検索キーワードに変換する。基本的にはデータベースのカテゴリと概念識別子が一致しているため、「ホテル」の「AAA3」のように概念識別子だけで条件が生成できる。しかし、「河口湖」のように「場所」のカテゴリと「施設名」としてのカテゴリがある場合は、一意に検索項目が決定できない。そこで、キーワード変換則にあらかじめ「goal」格で係っている場合は、場所で検索するというルールを用意しておくことにより、「場所」のカテゴリでの検索キーワードを生成する。これにより最終的に検索条件『(場所河口湖)(種類 ホテル)』の検索キーワードが生成できる。

### 3.3 実行結果

検索キーワードが得られたら、SQL 生成部にて SQL 構文を生成し、検索結果が得られる。図 5 に富士山観光案内システムにおける検索例を示す。

## 4 富士山観光案内システムでの評価

### 4.1 辞書の検索タスク適応

まず、富士山観光案内のドメインにおいて本稿の汎用的な意味理解部を用いてシステムを適用した。その際、従来システム [4] をデータベース検索タス

入力文: 河口湖にはどんなホテルがありますか

```
=== 意味解析結果 ===
*[ある. 動詞-自立]
object [ホテル. 名詞-一般]
goal   [河口湖. 名詞-固有名詞-一般]
      a-object [どんな. 連体詞]
```

```
=== 検索キーワード ===
(場所 河口湖) (種類 ホテル)
```

```
=== SQL 構文 ===
select distinct(dat.dat_id)
from dat,loc,kind
where ((loc.loc = '河口湖')
      and dat.dat_id = loc.dat_id)
      and ((kind.kind = 'ホテル')
          and dat.dat_id = kind.dat_id);
```

```
=== システム出力 ===
システム: 4 件見つかりました
```

```
=== 検索結果 ===
1 番. レイクサイドホテル (ホテル), 河口湖, 12000 円
2 番. レイクランドホテル (ホテル), 河口湖, 12000 円
3 番. 河口湖ホテル (ホテル), 河口湖, 12000 円
4 番. 河口湖グランドホテル (ホテル), 河口湖, 12000 円
```

```
=== システム出力 ===
システム: 検索結果は以上です
```

図 5: 検索結果表示例

クへの適応において、データベース検索システムの枠組みにあてはめて作成し直した。従来システムで、評価されている収集済みのユーザ発話文 499 文を利用して意味理解できるようなシステムを構築してきた。しかし、データベース検索というタスクに限定した場合、汎用的な意味辞書である EDR 電子化辞書では足りない単語や関係が頻出する。そこで、データベース検索システムとして必要な辞書をあらかじめ収集した発話例 499 文を参考にして補うことにした。

これらに沿って、解析できない単語、関係について以下のように辞書の追加を行った。

- 不足している格フレーム (日本語共起パターン副辞書) への登録  
例文を解析していて、形成できない格フレームのものについては随時登録を行った。ここでは 5 個の格フレームの追加または編集が必要となった。
- 不足している単語の登録  
利用したい単語であるのに、EDR 電子化辞書にないような単語 (“どのような: 連体詞” など)、または特別扱いの単語 (“です: 判定詞”) は扱いやすいように登録をした。その他、検索タスクでは 1 単語として扱いたい単語 (“教えて下さる: 動詞”) などを含め 9 単語を新規に登録した。それに伴い、新しい概念識別子のカテゴリを作成したことによる新概念を 2 つ追加した。

このようにドメイン・タスクへの適応を行った意味理解部の評価を行った。1 つめの評価として、ど

れだけ正しく意味理解できているかを評価した。また、データベース検索タスクを達成する上で、ユーザの発話文に対して、正しいキーワードを生成して、正しい検索ができたかという評価も行った。意味解析の成功率は82.5%、検索成功率は85.1%を達成することができた。以前のシステム [4] では意味解析率が79.4%で検索成功率が79.4%であったことから汎用的な理解部によって高精度で検索可能なレベルを達成できたと言える。

## 4.2 音声認識文による評価

今回新しく作成した意味理解部の評価を行うため、開発環境と同じドメイン・タスクである富士山観光案内に関する4名の被験者による発話文とその認識文、合計399文を用いて評価を行った(前節の499文とは異なる。単語認識率:91.5%)。音声認識文と書き起こし文でどれだけ正しく意味理解できているかを評価した。

また、タスクの達成を示す上で、質問に対する正しい検索ができたかをも合わせて評価した。それらの結果を表1に示す。

表 1: 富士山観光案内システムの意味解析率・検索成功率

種類	総数	意味解析	検索成功
音声認識文	399	342(80.2%)	342(85.7%)
書き起こし文	399	364(91.2%)	326(90.7%)

書き起こし文については、4.1節における開発環境での評価文での意味解析率、検索成功率とほぼ同程度の値が得られている。また音声認識文については、助詞落ちや文頭文末の誤認識の対応はできていないが、キーワードとなる単語が認識できていないと意味理解も検索も不可能である。それらが原因となって単語認識率に比べ解析できた割合が低くなってしまったという結果になった。

## 5 文献検索システムへの適用

富士山観光案内ドメインでは、タスクへの適応にも同ドメインの例文を参考にしたが、開発用と同じドメイン・タスクではなく、文献検索ドメインへの適応を行ったシステムでの評価を示し、本意味理解部の汎用性を実証する。

学術情報センターの文献検索システムに対する対話データ [12] と、電子情報通信学会論文誌、情報処理学会論文誌、日本音響学会誌、人工知能学会誌の日本語の論文情報(3161件)を実際に使用した。システムを適応するのに際して、富士山観光案内システムでも行ったようにデータベース単語を登録する必要がある。「著者」「著者姓」「著者名」「キーワード」の4つのカテゴリについて形態素解析辞書、意

味辞書に固有名詞を中心に16,920単語を登録した。これはツールによって登録情報を与えることによって自動登録が可能である。それ以外は開発用の辞書から手を加える必要はない。

実際に文献検索システムに適用したシステムの動作例を示す。「山岡克式さんの論文はありますか」と入力した場合の検索結果を図6に示す。入力文に対応した意味表現、それに対応した検索キーワード、SQL構文が生成され、検索結果が一覧で表示されている。

```

入力文: 山岡克式さんの論文はありますか

=== 意味解析結果 ===
*[ある. 動詞-自立]
object [論文. 名詞-一般] modifier |山岡克式. 名詞-固有名詞-人名-一般|[さん. 名詞-接尾-人名]

=== 検索キーワード ===
(著者 山岡克式)

=== SQL 構文 ===
select distinct(ref_ref_id)
from ref_auf
where ((auf.famlas = '山岡克式')
and ref_ref_id = auf.ref_id);

=== システム出力 ===
システム: 2件見つかりました

=== 検索結果 ===
1番: 酒井善則, 山岡克式, 菅原真司, 『エントロピーに基づいた情報探索に関する検討』, 電子情報通信学会論文誌. A. 基礎・境界, (1995)
2番: 山岡克式, 吉田俊之, 酒井善則, 『データユニット予測に基づくマルチメディア同期法』, 電子情報通信学会論文誌. B-1. 通信1-情報通信システム・理論, (1997)

=== システム出力 ===
システム: 検索結果は以上です

```

図 6: 実行例

Wizard of Oz 法による過去の評価実験 [4] で得られたテスト文100文のうち、データベース検索タスクに該当する34文を3名が発話した発話文に対し、富士山観光案内システムと同様に文献検索システムでの音声認識文(単語認識率:60.9%。言語モデル作成用のコーパスが小規模なため低い認識率)と書き起こし文でどれだけ正しく意味理解できているかを評価した。

また質問に対する正しい検索がどの程度できたかを評価した。その結果を表2に示す。

表 2: 文献検索システムの意味解析率・検索成功率

種類	総数	意味解析	検索成功
音声認識文	204文	160(78.4%)	189(92.6%)
書き起こし文	34文	33(97.1%)	33(97.1%)

発話文が「～さんの論文はありますか」、「～に関する論文を教えてください」のような決まった文型が多いこともあり、書き起こし文についてはほとんど

ど解析できている。音声認識文では単語認識率が約60%と高くはないにも関わらず、ロバストな意味理解ができていると考える。キーワードとなるような単語が得られていれば、意味理解、検索ともに認識率の低さをカバーできていると言える。また、文献検索タスクの単語が富士山観光案内のようにデータカテゴリを複数持つものがないことも大きな要因である。このように、文献検索のような富士山観光案内に比べて小さなものである場合、十分に適用可能である。

## 6 ホテル検索システムへの適用

文献検索システムのように極めて限られたドメイン・タスクにおけるものでなく、より複雑な対話を目指してホテル検索システムの構築を行った。ホテル検索・予約サイト HORNET<sup>2</sup>にあるホテルデータのうち、東京都内にあるデータ180件のホテルを対象に行った。検索カテゴリとしては、「住所」「アクセス」「シングル料金」「ダブル料金」「ツイン料金」「サービス」などの16項目から検索できる。これより680単語について各種タスク同様に形態素解析辞書、意味辞書に登録を行った。

ホテル検索システムに適用したシステムの動作例を示す。「東急線沿線にホテルはありますか」と入力した場合の検索結果を図7に示す。

```

入力：東急線沿線にホテルはありますか

=== 意味解析結果 ===
* [ある, 動詞-自立]
goal | 東急線, 名詞-固有有 一般 | [沿線, 名詞-一般]
object | [ホテル, 名詞-一般]

=== 検索キーワード ===
(鉄道会社 東急線)

=== SQL 構文 ===
select distinct(htl.htl_id)
from htl,acc
where ((acc.access like '%東急線%')
       and htl.htl_id = acc.htl_id);

=== システム出力 ===
システム：2件見つかりました

=== 検索結果 ===
1番 東横イン蒲田1, 大田区西蒲田8-7-7, JR 線蒲田
  駅から徒歩5分, 東急線 蒲田駅から徒歩5分, 羽田
  空港からバス35分, 京浜急行京急蒲田から徒歩13
  分, S 10000円, D 10000円, T 10000円
2番 東横イン蒲田2, 大田区西蒲田7-24-7, JR
  線蒲田駅から徒歩5分, 東急線 蒲田駅から徒歩5
  分, S 10000円, D 10000円, T 10000円

=== システム出力 ===
システム：検索結果は以上です

```

図7: 実行例

ホテル検索システムでは地名、鉄道路線名などが大量にあるため、音声認識用の十分な言語モデルを短期間で作成するのは困難である。そこで、より自

然な対話文を12名に考えてもらった文、および2名が電話形式でオペレータ役(人間)と対話して収集した文のうち、データベース検索システムの入力文に該当する文111文についてどれだけ正しく意味理解できているか、また質問に対する正しい検索ができたかを評価した。

表3: ホテル検索システムの意味解析率・検索成功率

種類	総数	意味解析	検索成功
ホテル検索用	111文	67(60.4%)	68(61.3%)

ホテル検索においては、複雑な対話文が多く単語数も検索のカテゴリ、バリエーションも多岐にわたるため、これまでのシステムに比べて意味理解、検索とも6割程度しかできないという結果になった。意味理解できなかった例を図8に示す。

```

・5000円くらいはありますか
  --形式名詞「の」に対応できず、概念識別子がない。

・京浜急行の駅がそばにあるホテルはありますか
  --格決定の際、「そば」の格が「蕎麦」の意味でとられ、
  "agent"格になってしまう。

・そのホテルのシングルとツインの料金教えてください
  --係り受け構文解析器(南瓜)で、シングルとツインの
  並列関係がとれていない。

```

図8: 実行例

同じ検索タスクではあるが、富士山観光案内システムで想定した発話文のレベルを越えるものには対応できなかったり、辞書の対応タスクに対し不十分で、タスクに適した格決定であるという点が大い。汎用的な辞書を限られたドメイン・タスクで利用するにはそれにあった辞書のカスタマイズがある程度必要であると言える。

意味理解部全体としては、関係を正しく得るためにはタスク適応による部分が大いと考えられる。EDR電子化辞書をオリジナルのまま使用して、本研究での意味理解手法を用いると深層格がそのタスクに合わない理解を示す場合も多かった。今回の意味解析手法では、深層格決定に関して概念距離の考えなどまだ不十分な部分もあると考えられるし、使用している辞書(日本語共起パターン副辞書、概念記述辞書)のみで十分であるとも言えない。

また、データベース用の単語を登録する際に、ユーザの主観によっておおよその概念体系の位置に登録しているが、これも意味理解部に大きく影響を与えるため、慎重に行う必要がある。

さらに、形態素解析、係り受け構文解析には汎用ツールを使っているため、それらで誤りやツールの限界を越える処理、またはタスクにそぐわない解析が行われると意味理解部では解析ができないということになってしまう。

<sup>2</sup><http://www.inn-info.co.jp/>

## 7 移植コストの評価

既存の富士山観光案内システムから文献検索システム、ホテル検索システムに移植する際の意味理解部の移植性について調べた。すでにデータベースが用意されている状態から、既存の茶釜辞書、意味辞書にデータベース検索のための固有名詞等を登録する作業を行う。その際、データベースを構築する際に得られるカテゴリごとのデータベース登録単語一覧より、品詞やEDR電子化辞書の概念体系に組み込む際の上位概念子を指定することで、すべての単語がツールによって自動的に登録することができる。この作業を本音声対話システムの開発に携わったことのない大学院生によって行ったところ、文献検索システムでは4カテゴリ16,920単語を登録するのに約1.5時間・人で登録が可能であった。従来の意味理解部[4]では、文献検索システムにおいて形態素解析辞書の作成に2時間・人、意味辞書の作成に10時間・人を要していたことを考えると比較すると非常に移植性が改善されたと言える。

しかし、検索キーワード生成に関してのキーワード変換則の生成には簡単なルールのみであれば簡単に用意できるが、形容詞による複雑な条件制約や複雑なルールについては本意味理解部開発者に尋ねないと実現できなかった。

また、カテゴリが多いホテル検索システムにおいても同一人物によって移植を行った。登録情報が多いにも関わらず、登録作業に慣れてきたため、10カテゴリ680単語を登録するのに1.5時間・人程度で登録が可能であった。

## 8 まとめ

汎用性の高い音声対話システムのために意味理解部を汎用的にするための改良を行った。そのため、意味理解部と検索部による汎用化を汎用的な日本語辞書であるEDR電子化辞書の導入を中心に開発を行った。

EDR電子化辞書を利用した意味理解に対応すべく形態素解析、係り受け構文解析ツールの導入から、意味理解手法の手順までの流れ、細かい意味理解手法にわたるまでを明確にし、検索結果を得るまでの手法を示した。また、音声対話システムに対応すべく音声入力を前提としたロバスタな意味理解ができる処理も組み込んだ。

また、データベース検索タスクにおいてさまざまなドメイン・タスクにおいてEDR電子化辞書を用いて移植性のコスト削減を目指した。少ないドメイン・タスク情報でさまざまなドメイン・タスクへの移植が可能であることを「富士山観光案内」、「文献検索」、「ホテル検索」の3つにおいて示した。それらの意味理解率、検索成功率においても一定の評価を得ることができた。

このシステムではデータベース検索タスクのみしか実現していない。今後は、意味理解部の汎用性を生かして、その他のタスクでの実現や的確な応答生成が望まれる。また、意味理解部はドメイン・タ

クに依存していない点を生かして、ドメイン・タスクに関係ない発話への対応ができたり、さらに対話中に話題が変わった場合のドメイン・タスクへの変更が用意になったりするなどの発展が考えられる。

## 参考文献

- [1] Denecke, M.: Rapid Prototyping for Spoken Dialogue System, Proc. COLLING2002 (2002.8).
- [2] Nakagawa, S., Kogure, S. and Itoh, T.: A Semantic Interpreter and a Cooperative Response Generator for a Robust Spoken Dialogue System, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 14, No. 5, pp. 553-569 (2000).
- [3] 小暮悟, 伊藤敏彦, 中川聖一: 音声対話システムの移植性に関する考察—観光案内システムとデータベース検索システム—, 情報処理学会研究報告, 99-SLP-25, pp. 13-18 (1999.2).
- [4] 小暮悟, 中川聖一: データベース検索用音声対話システムにおける移植性の高い意味理解部・検索部の構築と評価, 情報処理学会論文誌, Vol. 43, No. 3, pp. 714-733 (2002.3).
- [5] 原田実, 田淵和幸, 大野博之: 日本語意味解析システム SAGE の高速化・高精度化とコーパスによる精度評価, 情報処理学会論文誌, Vol. 43, No. 9, pp. 2894-2902 (2002.9).
- [6] Umeda, M., Kogure, S. and Nakagawa, S.: Interpreter for Highly Portable Spoken Dialogue System, Proc. SIGdial Workshop on Discourse and Dialogue, pp. 105-114 (2003.7).
- [7] (株) 日本電子化辞書研究所: EDR 電子化辞書使用説明書 (2001.8).
- [8] 奈良先端科学技術大学院大学: 日本語形態素解析システム 茶釜 version 2.3.3 (2003.8). <http://chasen.aist-nara.ac.jp/>.
- [9] 工藤拓, 松本裕治: チャンキングの段階適用による日本語係り受け解析, 情報処理学会論文誌, Vol. 43, No. 6, pp. 1834-1842 (2002.6).
- [10] Li, Y., Bandar, Z. A. and McLean, D.: An Approach for Measuring Semantic Similarity between Words Using Multiple Information Sources, IEEE Transactions on Knowledge and Data Engineering, Vol. 15, No. 4 (2003).
- [11] 田中徳積: 自然言語処理 -基礎と応用-, 電気情報通信学会 (1999).
- [12] 藤崎博也, 大野澄雄, 飯島岐勇: 音声対話の収録方法とその設定について, 学振未来開拓事業プロジェクト 音声対話資料収録用参考資料 (1998.1).