

文中の語義の説明を行う
対話的文書読み上げシステムの研究

細野 祥代

名古屋大学 工学部 電気電子・情報工学科

2003年2月

目次

第1章	はじめに	3
1.1	研究の背景と目的	3
1.2	本研究で提案するシステムの概要	3
1.3	本論文の構成	4
第2章	対話的文書読み上げシステム	5
2.1	システムの機能と構成	5
2.2	音声認識	5
2.3	音声合成	5
2.4	国語辞書データベース	6
2.5	使用する辞書	6
2.6	対話システム	9
第3章	文書の読み上げ	11
3.1	文書の選択	11
3.2	文書の読み上げ	11
3.3	同期の必要性	11
3.4	辞書引きする語の特定	12
第4章	辞書引きと語義の絞り込み	15
4.1	辞書引きの手順	15
4.2	語義の絞り込み	16
4.2.1	文法上の性質による絞り込み	17
4.2.2	前後の語の語義を利用した絞り込み	17
4.2.3	対話による絞り込み	17
4.2.4	複合語の場合	18
4.3	代表語	18
4.3.1	代表語の抽出	18
4.3.2	代表語の提示	19
4.3.3	ユーザからの代表語の入力	19
4.4	語義の説明文の生成	20
4.5	履歴情報の利用	20

第5章	関連研究	23
5.1	ダイアログナビ (Dialog Navigator)	23
5.2	会話インタフェースロボットとの状況依存会話	23
5.3	ATIS (Air Travel Information System)	23
5.4	Talkman	24
5.5	飛遊夢	24
第6章	おわりに	25
6.1	まとめ	25
6.2	今後の課題	25
6.2.1	認識辞書の動的作成	25
6.2.2	同義語・上位語の判定	25
6.2.3	文脈の考慮	26

概要

自然言語による対話は、人間にとって最も頻繁に使用されるコミュニケーション手段と言える。人は対話によって自分の意図を相手に伝えたり、相手と共同で問題の解決を図る。そこで、人とシステムの間でも対話によって問題の解決を図れないかと考えるのは自然である。本研究では、人間による文書の理解という問題を、対話によって支援するシステムを作成した。

文書を理解する上で、まず必要となるのは単語の意味を知ることである。文中に意味のわからない単語が出てきた場合、辞書を引く必要が出てくる。現在、辞書引きは単語単位で行われており、その後、複数の語義の候補の中からユーザが適切だと思うものを選択している。しかし、文書に含まれる単語の意味が知りたい場合、語義の選択等で内容の理解が妨げられるのは適切ではない。そこで、本研究では、語義の検索と絞り込みを、ユーザとの対話や意味的に構造化された辞書を利用しながら自動的に行う仕組みを実現した。また、複合語は頻繁に使用されるが辞書には記載されていないことが多い。そこで、複合語の意味をユーザから質問された場合には、要素となる単語の語義の説明文から適切な説明を生成して答えることで、より自然な対話を可能にした。

第1章 はじめに

1.1 研究の背景と目的

対話とは、人間にとって最も自然で慣れ親しんだコミュニケーション手段である。人は対話を行うことによって自分の意見や意思を相手に伝えたり、相手の意見や意思を理解することで、共同で問題の解決を図る。そこで、人と情報処理システムの間でも対話によって意思の疎通を図り、問題の解決を目指すことはできないかと考えるのは自然である。

近年、音声認識や音声合成などの重要な音声処理技術が盛んに研究・開発され、実用に耐えられるようになってきた。実際、カーナビゲーションシステムや携帯電話の音声認識機能など、日常生活に浸透しつつある。カーナビゲーションシステムでは、音声によるガイド、音声認識を利用したコマンド入力に検索、さらに声質の自動学習機能を持つものが登場してきている。携帯電話でも音声コマンド機能を持つものや、音声によってインターネットの情報を入手するサービスなどが実用化されている。また、形態素解析などの自然言語処理の要素となる様々な技術も成功を収めつつある。これらの技術に支えられて、電話の自動応答システムなどの音声対話システムに関する研究も盛んに行われている。

1.2 本研究で提案するシステムの概要

本研究では、ユーザとの対話によって文書の理解を助ける音声対話システムを作成した。文書を理解する上で、まず必要となるのは語の意味を知ることである。文中に意味のわからない単語が出てきた場合、辞書を引く必要が出てくる。現在、辞書引きは単語単位で行われており、その後、複数の語義の候補の中からユーザが適切だと思うものを選択している。しかし、文書に含まれる単語の意味が知りたい場合、語義の選択等で内容の理解が妨げられるのは適切ではない。そこで、本研究では語義の検索と絞り込みをユーザとの対話や意味的に構造化された辞書を利用しながら自動的に行う対話システムの実現をめざした。

また、複合語は頻繁に使用されるが辞書には記載されていないことが多い。このような場合には、複合語を要素となる単語ごとに分解し、単語ごとに辞書引きや語義の絞り込みを行い、得られた要素となる単語の語義の説明文から複合語の語義を推論するという煩雑な手順を踏んでいる。そこで複合語の意味をユーザから質問された場合には、システムが自動的に要素となる単語に分解、辞書引きや絞り込みを行い、語義の説明文から適切な説明を生成して答えることで、ユーザの負担を軽くすることをめざした。また、説明文から適切な説明を生成することによって、より自然な対話を可能にした。

1.3 本論文の構成

第2章では、本研究で作成したシステムの構成などについて述べる。第3章では、文書の読み上げの手順や仕組み、ユーザの割り込みについて述べる。第4章では、辞書引きの手順や仕組み、語義の絞り込みの方法や、語義を説明する応答文の生成について説明する。第5章では関連研究、最後の第6章では本研究のまとめと今後の課題について述べる。

第2章 対話的文書読み上げシステム

本章では、対話的文書読み上げシステムの機能と構成、その構成要素について説明する。また、文書の読み上げに関しては3章で、辞書引きと語義の絞り込みに関しては4章で詳しく説明する。

2.1 システムの機能と構成

システムは参考文献 [3, 2, 1] を参考にしながら構成した。対話的文書読み上げシステムは、Web 上のニュースなどの文書を読み上げながら、途中ユーザからの質問があった場合には読み上げを中断し、語義の説明を行うシステムである。語の意味の質問だけでなく、「最初から文書を読み直す」といった、読み上げに関する操作も発話によって行うことができる。システムは文書の読み上げ中にユーザからの呼びかけがあった場合、読み上げを一時停止してユーザからの要求を待つ。ユーザからの入力があると、システムは音声認識を行いユーザの発話の意味を特定し、それに基づいてユーザに応答する。また、システムは意味的に構造化された国語辞書データベースを持っており、ユーザから語義の説明を求められた場合には、このデータベースを検索し、適切な語義の定義文を返す。

対話的文書読み上げシステムは、図 2.1 に示されるように、音声認識、音声合成、対話システム、国語辞書データベース、処理の対象である Web 文書から構成されている。

2.2 音声認識

音声認識には IBM の ViaVoice を使用している。ViaVoice は Java Speech API(JSAPI) [4] に基づいて作られた Speech for Java を用いることで、Java のプログラム中から使用できる。音声認識部分では、認識辞書を用いてユーザの発話の意味を取得、対話システム部分にその意味をわたす。また、認識辞書は複数あり、状況に応じて使いわけている。

2.3 音声合成

音声合成には IBM の ProTalker97 を使用している。ProTalker97 は Microsoft Speech API に対応した TTS(Text-To-Speech) エンジン・オブジェクトを提供する。対話システムで生成されたユーザへの応答文、または読み上げの対象のテキストを対話システムから受け取り、音声合成を行う。

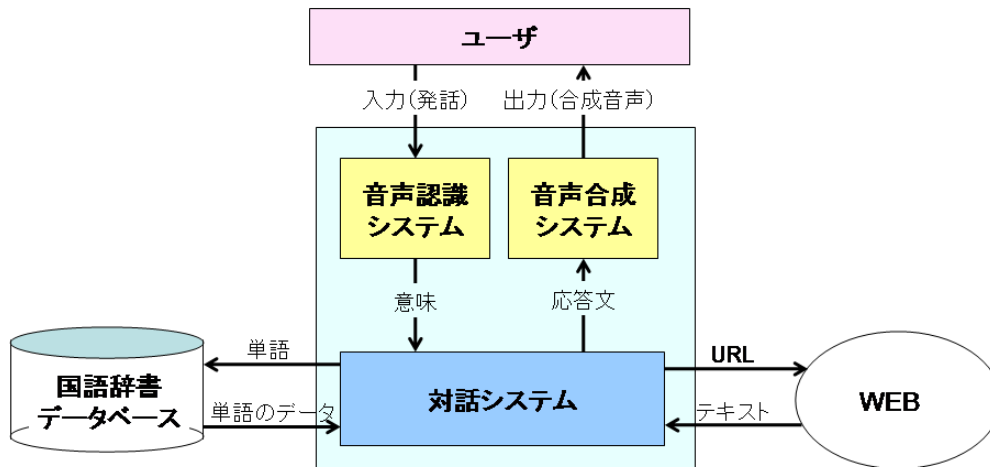


図 2.1: システムの構成図

2.4 国語辞書データベース

国語辞典データベースは、リレーショナルデータベースである PostgreSQL と XML データベースである Xindice [6] を併用している。岩波国語辞典には約六万二千語が収録されている。この国語辞典データベースのすべての語の中から見出し語などを検索する場合には、高速で検索を行うことが可能な PostgreSQL を使用する。ある語の語義の説明文といった限られた範囲内を検索するときには、Xindice を使用する。Xindice は問い合わせに XPath を用いている。XPath はパスのような書式で XML 文書の特定の部分を示すことができるため、XML の構造を生かした検索ができる。

2.5 使用する辞書

使用する辞書は岩波国語辞典を XML 化したもので、語義の説明文は構造化されている。また語義の説明文中にでてくる単語には、読み・基本形・品詞名・活用形・その他補足事項がアノテーション付けされている。これらの情報を利用しながら辞書引きを行う

XML 形式化された国語辞典の項目は以下のような構造になっている。

```
<?xml version="1.0" encoding="Shift_JIS" ?>
<entry entry="うんそう 運送 (表記)" resp="***** *****">
  <hd>うんそう (見出し) </hd>
  <sense id="04189.0" (語の ID 番号。語義がひとつの場合は語義の ID でもある。) >
    <sense type="1-1" id="04189.0.1" (語義の ID) >
      <orth>運送</orth>
      <pos>名・ス他 (品詞名や活用など。他にも接頭・接尾といった文法上の性質が記録されている。) </pos>
```

```

<su id="id0" ( su タグは一文ごとに区切る。id は説明文の番号。) >
  <vp ( v は動詞、 p はこの語が次に出てくるタグの語に係る語であることを示す。
  この場合は n タグの語「こと」に係る。) >
    <adp>
      <np syn="p">
        <np ( n は名詞) 読み="リヨカク" 品詞="名詞" sense="54696.0" ( 語義
  の ID ) >旅客</np>
        <seg ( 語以外 ) >・</seg>
        <np 読み="カモツ" 品詞="名詞" sense="09073.0.2">貨物</np>
      </np>
      <ad 読み="ヲ" 品詞="格助詞" sense="56250.1.1">を</ad>
    </adp>
    <v 読み="ハコブ" 品詞="動詞" 活用="五段・バ行-見出し形" sense="41062.0.1">
  運ぶ</v>
  </vp>
  <n 読み="コト" 品詞="名詞" 補足="非自立" sense="17814.1.3.2.4">こと</n>
  。
</su>
<eg> ( 用例文であることを示す )
「
  <su id="id1">
    <np ed="hd" 読み="ウンソウ" 品詞="名詞" sense="04189.0.1">運送</np>
    <n 読み="ギョウ" 品詞="名詞" sense="11800.0.3">業</n>
  </su>
  」
</eg>
</sense>
<sense type="1-2" id="04189.0.2">
  :
  :
</sense>
</sense>
</entry>

```

図 2.2 は岩波国語辞典を XML 化したファイルの例である。一つの語につきこのようなデータが一つ与えられる。

PostgreSQL での検索には、entry と hd の二つのタグの内容を使う。sense タグが複数ある場合、その語は多義語であり、一番内側にある sense タグ一つの下で一つの語義が説明されている。

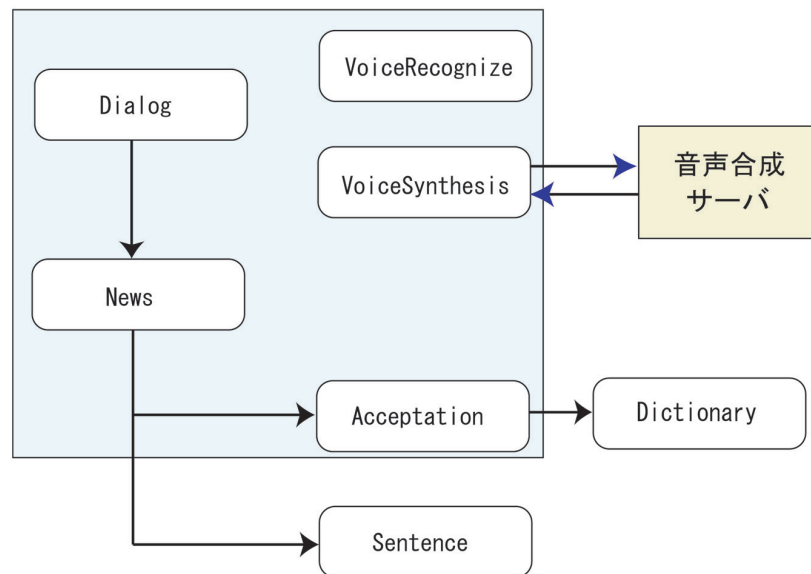


図 2.3: クラスの関係図

2.6 対話システム

対話システムでは、音声認識部分からわたされたユーザの発話の意味をもとにユーザへの応答を生成したり、プログラムの実行を行う。例えば、ユーザからの要求に応じて、ニュースを読み上げたり、国語辞書を引いたりする。

図 2.3 は、対話システムのクラスの関係を示した図である。
実装したクラスについて説明する。

1. Dialog

システムを起動し、対話を始めるためのクラス。保持する主な関数。

- (a) start
対話を始める。
- (b) selectProgram
ユーザからの入力に応じて、ニュースを読んだり、システムを終了する。
- (c) end
システムを終了する。

2. News

ニュースを読むためのクラス。保持する主な関数。

- (a) selectNews
ユーザとの対話によって読み上げるニュースを選ぶ。
- (b) read
ニュース記事を読み上げる。
- (c) seekUnknownWord
ユーザが意味のわからない語を、文章中から探す。

3. Acceptation

ユーザの意味の分からない語の語義を特定、取得する。保持する主な関数。

- (a) narrowDown
国語辞書データベースを検索した結果、語・語義が複数あった場合に絞り込む関数。
- (b) makeKeyWord
語義の説明文中からキーワードの抽出を行い、得られたキーワードのリストを返す。

4. Dictionary

辞書引きしたい語の表記、読みをもとに辞書引きを行う。保持する主な関数。

- (a) Dictionary
表記、または表記と読みをもとに国語辞書データベースを検索、一致した語のIDを保持する。
- (b) getAllSenseNode
検索した結果の語のすべての sense ノードを配列にして返す。
- (c) getSenseByID
複数の語義から一つの語義を取り出す際に呼び出される。指定されたIDを持つ sense ノードを返す。

5. VoiceSynthesis

音声合成サーバとの通信を行う。保持する主な関数。

- (a) speak
音声合成サーバに発話させたい文字列を送信する。

6. VoiceRecognize

ViaVoice を利用して音声認識するためのクラス。保持する主な関数。

- (a) listen
ユーザの発話を音声認識する。また、音声認識に使用する認識辞書を指定する。

第3章 文書の読み上げ

本章では、文書の読み上げと、それに関する対話について説明する。

3.1 文書の選択

システムが文書を読み始める前に、読み上げる文書を選ぶ必要がある。本システムでは、いくつかのジャンルのニュース記事からユーザの興味のあるものを選んでもらう。表 3.1 は、ユーザとシステムの対話によって読み上げるニュースを選ぶ対話例である。

3.2 文書の読み上げ

本システムでは、対話システム側がクライアント、音声合成側がサーバとなっている。対話システム側では、まず、読み上げる文書を一文単位に分割する。一文単位で一度分割するのは、文書を途中から読み直す際に文頭から読み直した方が理解しやすいと考えたからである。その後、句点、鉤括弧でさらにこまかく分割する。それから、分割されたものを順に音声合成部分に送る。このとき、音声合成（サーバ）と対話システム（クライアント）間で同期をとることで、送信した単語列の音声合成が終わるまでは対話システム（クライアント）は次の単語列を送らないようにしている。これは、文書の読み上げ中に、ユーザからの割り込み（発話）があった場合に、テキストのどの部分を読み上げていたのかを知るためである。

3.3 同期の必要性

読み上げていた位置を知る必要があるのは、以下のような理由による。

一つ目の理由は、読み飛ばしの防止である。同期をとらずに次々と文を音声合成部分に送り続けると、音声合成が間に合わなくなり読み飛ばされてしまうことがあった。この読み飛ばしを防ぐために、前の文を読み終わってから次の文を音声合成部分に送るようするため同期をとる必要があった。

二つ目の理由は、ユーザからの「今のところもう一回読んでほしい」といった読み上げる位置に関する要求に答えるためである。

そして最後の理由は、ユーザがどの語の意味が分からないのかをシステムが知るのに、読み上げていた位置の情報を使用するためである。

表 3.1: ニュースの選択

話者	発話
system	ななに
user	ニュースを読んで
system	どんなニュースが良い?
user	どんなのがあるの?
system	国内、海外、経済、コンピュータ、エンターテイメントがあるよ。
user	経済がいい。
system	全日空が不採算路線を整理、...
:	:
:	:

3.4 辞書引きする語の特定

辞書引きする語の特定の仕方は次のようになっている。

読み上げられる文書中にわからない語があった場合、ユーザはそのことをシステムに伝えたい。例えば、ユーザはある文書中の「N社が不採算路線から撤退」という部分の「撤退」という言葉がわからなかったとする。ここで、ユーザは「撤退という言葉の意味がわからない」とシステムに伝えたい。しかし、「撤退という言葉の意味がわからない」という発話を ViaVoice で音声認識させるためには、認識辞書に「撤退」という言葉を登録しておかなければならない。システムには、ユーザにどの語の意味がわからないと言われるかはわからない。そのため、「～という言葉の意味がわからない」という発話に対応するには、文書中に出てくるすべての単語をあらかじめ認識辞書に登録しなければならない。しかし、認識辞書に登録できる語数には限界があり、また文書中に出てくる語は数十から、ものによっては数百、またはそれ以上あることもある。そのため、全ての語をあらかじめ認識辞書に登録することは困難である。また、あらかじめ認識辞書に全ての語を登録できたとしても、他の文書を読みあげるときには再び新しく認識辞書を作り直さなければならない。

そこで今回は、意味のわからない語があった場合、ユーザは「意味のわからない語がある」ということをシステムに告げることにした。その後、システムはユーザとの対話によって、文中のどの語の意味がわからないのかを特定する。ユーザは理解できない語が出てきた時点で、システムに意味のわからない語があったことを告げることが多いと思われる。そこで、システムはユーザからの発話があった時点で読み上げていた部分から、名詞・動詞・形容詞・形容動詞・副詞を取り出し、順にユーザに提示していく。複合名詞の場合は、その要素となる名詞に分解せずに複合名詞のままでユーザに提示する。これは、ユーザの側では複合名詞をひとつの語としてとらえていることが多いと考えたからである。助詞・助動詞は提示しない。なぜなら、助詞・助動詞の意味をきかれる回数は、名詞などの意味を質問される回数よりも少ないと思われるからである。また、助詞や助動詞だけでは意味をなさないことから質問されることは少ないと考えた。

表 3.2 はユーザはどの語の意味がわからないのかを特定するための会話の例である。「ちよっ

表 3.2: 割り込みと語の特定

話者	発話
system	... (ニュースを読み上げ中) ...
system	...価格競争が厳しくなっていることから、
system	国内便 5 路線から撤退、
user	ちょっと待って
system	なあに？
user	わからない言葉があるんだけど
system	えー
system	国内便 5 路線がわからないの？
user	違う
system	撤退がわからないの？
user	そう
system	撤退だねー
：	：
：	：

と待って」というのがユーザの割り込みで、その後、システムが割り込みがあった付近の語を順に提示することで意味がわからない語を特定している。

第4章 辞書引きと語義の絞り込み

対話的文書読み上げシステムでは、ユーザは文書の読み上げ中に意味のわからない語があった場合、システムに読み上げを一時停止して語の意味の説明を求めることができる。本章では、この辞書引きと語義の絞り込みについて説明する。

4.1 辞書引きの手順

ユーザが意味を知りたい語が特定できたら、システムは次の手順でその語の辞書引きを行う。

1. 形態素解析

辞書引きの対象となる語の形態素解析を行う。形態素解析には茶筌 [5] を使用している。形態素解析を行うと、その語の表層語・基本形・読み・品詞・活用をえることができる。複合語の場合は、形態素解析を行うと要素となる単語ごとに分割される。

2. 国語辞書データベースの検索

基本形と読みを使い、Xindice で国語辞書データベースを検索する。対象となる語が複合語ではなかった場合には、基本形と読みをそのまま使用すればよい。複合語の場合は、形態素解析を行うと要素となる単語ごとに分割されるが、まずは分割された語を元通りに連結し、複合語の形で国語辞書データベースを検索する。これは、例えば「代金引換」などのように複合語でも辞書に登録されている場合があるためである。

また検索した結果、該当する語が辞書中になかった場合で、かつ辞書引きの対象となっている語が複合語ではなかったときには、辞書中にその語が載っていない旨をユーザに伝え、検索を終了する。複合語であった場合には、今度は要素となる単語ごとに国語辞書データベースを検索する。

3. 絞り込み

検索した結果、語義が複数あった場合には語義の絞り込みを行う。絞り込みについては次章で詳しく述べる。

4. 応答文の生成

検索した結果、語義の個数が一つだけであった場合、または絞り込みによって語義を特定できた場合、その語義をもとにユーザに語の意味を説明する応答文を生成する。辞書に載っていない複合語の場合、要素となる単語の語義の説明文から適切な説明文を生成する。辞書に載っていない複合語の語義の説明文の生成の仕方は***で述べる。

表 4.1: 語義の説明

話者	発話
system	...兵庫県警は24日までに、
system	N ハムで食肉部門を統括していた...
user	待って
system	なあに
user	わからないんだけど
system	えー
system	N ハムがわからないの？
user	違う
system	食肉部門がわからないの？
user	違う
system	統括がわからないの？
user	そう
system	統括だねー
system	別々になっているいくつかのことを、
system	まとめて一つにくくることだって。
system	続きを読んでも良い？
user	うん
:	:
:	:

表 4.1 は、文章中の語（図中では「統括」）の語義を説明している例である。

4.2 語義の絞り込み

語の中には複数の語義を持つものも少なくはない。辞書に記載されている複数の語義をそのまま全てユーザに伝えるのでは、ユーザにとって負担である。そこで、読み上げ中の文書、辞書や対話を利用して語義の絞り込みを行う。語義の絞り込みは以下のような段階で行われる。

1. 文法上の性質を利用した絞り込み
2. 前後の語の語義を利用した絞り込み
3. 対話による絞り込み

表 4.2: 「運動」の語義の一覧

表記	番号	語義
運動	1	物体（質点）が時の経過と共に位置を移す物理現象。
	2	健康や楽しみのために、体を動かすわざ、特にスポーツ。
	3	目的達成のために、いろいろな方面に働きかけて努力すること。

4.2.1 文法上の性質による絞り込み

接頭語や接尾語などといった文法による語義の絞込みを行う。例えば、「現実的」という言葉を形態素解析にかけると「的」は接尾語であると帰ってくる。そこで「的」の語義の上に「接尾語」というタグを持つ語義のみに絞ることができる。

4.2.2 前後の語の語義を利用した絞込み

前後の単語の語義からその語義の代表となる語（上位語、関連語等。以下、代表語と呼ぶ）を取り出し、それを利用して語義を絞り込む。代表語は語義を絞り込みたい語のより近くに存在する語のものから使用する。まず語義を絞り込みたい語の一つめの語義の中に、その代表語が含まれているかどうか調べる。含まれていた場合には、その語義を採用する。含まれていなかった場合には、次の語義について同様にして調べる。

例を挙げて説明する。「選手代表」という複合語は「選手」と「代表」という二つの単語から成る。「選手」の語義は「（競技に出るために）選ばれた人」というもの一つである。この語義の中から代表語の候補として「競技」「出る」「選ばれた」「人」が抽出される。これらに「選手」を加えたものを代表語とする。「選手代表」の語義の中で、それらの代表語のいずれかを含む語義を探すと「団体や～表すこと。また、その人。」という語義が見つかるので、これを採用する。

4.2.3 対話による絞り込み

上記2つの操作を行っても語義を特定できない場合は、ユーザとの対話によって語義を絞り込む。具体的には、語義ごとに代表語を抽出してユーザに提示することで語義を選択してもらう。

例えば「運動」という語を国語辞書で調べると、表 4.2 にあるとおり3個の語義がある。

そこで、システムはまず1番目の語義「物体（質点）が時の経過と共に位置を移す物理現象」から「物体」「現象」などを代表語として取り出す。次に2番目の語義から「健康」「スポーツ」などの代表語を取り出す。これを最後の語義まで繰り返して、全ての語義から代表語を取り出す。

その後ユーザに代表語を提示して一番近いと思うものを選んでもらう。ユーザに対する提示の仕方は、「1番、物体、現象。2番、...」と、全ての語義の代表語を順に伝える。ユーザは、システムの発話を聞き、何番が一番近いと思うかを告げる。表 4.3 は、実際に運動

表 4.3: 語義の絞り込み

話者	発話
system	運動って言葉なんだけど、
system	次の3個のどれに関係あるか選んでよ
system	1番：物体、現象
system	2番：健康、スポーツ
system	3番：達成、努力
user	3番
system	3番だね
system	運動の意味はね、
system	目的達成のために、
system	いろいろな方面に働きかけて努力することだって。
⋮	⋮
⋮	⋮

の意味を対話によって絞り込んでいる例である。

ユーザは、システムの発話を最後まで聞いてから選びたいときもあれば、最後まで聞かなくても、自分が知りたい語義がどれなのかわかるときもある。そこで、システムが発話の途中でも、該当するものを発見したときには割り込むことを可能にした。

4.2.4 複合語の場合

複合語の場合には、二つ以上の単語が語義を複数持つ場合がある。その場合は、どれか一つの語の語義が決定した場合、その語義の説明文からも新たに代表語を取り出し、残っている単語の語義の決定に利用する。

4.3 代表語

前章で、語義の説明文中の代表となる語（関連語、同義語、上位語等）を代表語と呼ぶことにした。ここでは、代表語の抽出、提示、入力について説明する。

4.3.1 代表語の抽出

代表語（上位語、関連語等）はシソーラスは利用せずに、国語辞書の語義の説明文から取り出している。代表語の生成にシソーラスを用いない理由は、シソーラスは発想を広げるには役立つが厳密に絞っていくことは難しいと考えたためである。

代表語の取り出し方としては、まず「日本語では重要な語は文の後のほうに出てくると多い」というヒューリスティックな規則に基づき、最後に出てくる単語に注目する。実

際、国語辞書の語義の説明文では上位語は最後にあることが多い。また文末には上位語ではなくても、より噛み砕いた同義語が出てくることも多い。ただし、「こと」「もの」「すること」といった単語が最後にある場合にはその上にある単語を見なければならない。

次に、文頭に注目する。文頭にも関連語が出てくることは多い。日本語では、修飾語が前で非修飾語が後に来ることが多い。しかし、国語辞書の語義の説明文は修飾の少ない簡潔で短い文で書かれているため、関連語が文頭付近にあることは多い。

例えば、「運航」という単語の語義は「船・航空機が決まった航路を動いていくこと。」であるが、「船」「航空機」は「運航」という言葉の関連語と言える。語義の説明文では、動作などの主体となる語や目的語、具体例が文頭に来ることが多い。これら主体となる語や、目的語は辞書引きの対象となる語に密接に関係していることが多いため、代表語と成り得ると考えた。また、上位語では漠然としすぎてしまう場合には、上位語よりもむしろ関連語の方が語義の選択に役立つことがあると考えた。「運航」という語では、「動く」という語が上位語になるが、ユーザに対して「動く」に関係があるかと聞くよりも、「船」航空機」に関係があるかと聞いた方が、ユーザには分かりやすいと思われる。

そこで、代表語の抽出は、語義の説明文の前後から同時に見ていくことにした。最初に語義の説明文中の単語から、属性の品詞が名詞・動詞・形容詞・形容動詞という代表語候補となる語をとりだす。まずは一番最後の語から調べる。このとき、調べた語が「こと」「もの」などであった場合にはその一つ前の語を調べる。上記の作業を繰り返して、後ろから語を順に調べていき、国語辞書中で語義の説明がされているものであれば代表語として採用する。次に代表語候補の中の先頭の語を調べる。先頭の語も、同様に国語辞書中で語義の説明がされているものであれば代表語として採用する。以後同様にして順に説明文の中心部分に向かって単語を調べ続けていく。

4.3.2 代表語の提示

ユーザに提示する代表語は一語か二語である。語義の説明文中の単語を調べて、二つ以上の代表語が見つかった場合には二つ提示する。二語提示するのは、上位語では漠然としている場合などに、関連語で補えないかと考えたためである。ユーザに提示する代表語の数を二個までとしたのは、音声対話の場合、システム側の発話があまりに長くなると、聞いていてユーザにストレスを感じさせてしまうと考えたためである。またユーザに提示する代表語は二語までであるが、システム内での語義の絞り込みには残りの代表語も使用している。また、システム側から提示された代表語の語義をさらに質問することもできる。システム側から提示される代表語は語義の説明文中の単語であり、属性としてその文に適した語義のIDを持っている。そのため、代表語の語義を質問されても語義の絞込みの必要はない。

4.3.3 ユーザからの代表語の入力

対話による絞り込みの段階では、システム側からの代表語の提示だけではなく、ユーザからも代表語を提示することが可能である。ただし、ユーザから入力できる代表語は、あらかじめ認識辞書に登録してあるものしか入力できない。

4.4 語義の説明文の生成

単語、もしくは国語辞書中で語義の説明のされている複合語の場合は、その語義の説明文を利用する。説明文そのままでは、対話として自然な文ではない。そこで、語尾に付け足すことでより口語的な説明文に直している。本システムでは辞書中の説明文が名詞で終わっている場合には、語尾に「だって」をつけている。動詞で終わっている場合には、語尾に「らしいよ」をつける。形容詞で終わっている場合には、語尾に「ことだって」をつける。助動詞で終わっている場合には、語尾に「ってことだよ」をつけている。例えば、「採算」の語義の説明文は「収支を計算してみること。」である。説明文は名詞で終わっているので、語尾に「だって」を付け足した、「収支を計算してみることだって。」という文でユーザへ応答文とする。

そのままの形では、国語辞書中で語義の説明をされていない複合語の意味を質問された場合、まず要素である単語ごとの語義を決定する。全ての要素の語義が決定すると、その語義の説明文から適切な一文を生成し、ユーザに応答する。要素である単語ごとに語義を説明するのではなく、複合語全体でひとつの説明文を返すのは、その方が対話として自然だと考えたためである。説明文の生成の方法は、語義の説明文中にある、主語・目的語部分を前の語の語義と置き換えるというものである。例えば、「保存運動」という語の場合、「保存」「運動」の2語の語義から説明文を生成する。それぞれの語義は「そのままの状態を保つようにして、とっておくこと」、「目的達成のために、いろいろな方面に働きかけて努力すること」である。「目標達成」部分を「保存」の語義と置き換えて、「そのままの状態を保つようにして、とっておくことのために、いろいろな方面に働きかけて努力すること。」をユーザへの応答文とする。

しかし、複数の単語の語義からひとつの説明文を自動的に生成することができないこともある。単語どうしの関係がわからないために単語の語義は決定していても複数の意味に解釈できてしまうためである。このような場合には、一つの文にはせずに単語ごとに語義を説明することにした。間違った意味をユーザに説明することは、文書の理解をかえって妨げてしまうからである。

4.5 履歴情報の利用

履歴情報には、システムが実行した機能が実行した順に記録されている。記録されるのは、どの文書を読み上げたのか、ユーザがどんな単語の辞書引きをしたのか、また辞書引きした単語の語義などである。これを利用して複合語中の単語で以前説明した単語の部分の語義の説明を繰り返ささないようにしている。例えば、ユーザが「不起訴処分」という単語の意味を質問した後で、今度は「不起訴釈放」という単語の意味を質問してきた場合、「不起訴」の語義は同じであるので「不起訴」はそのまま「釈放」の語義と結合してユーザに返す。一つの単語が複数の語義を持っていることがあるため、履歴情報を参照するのは、すべての語の語義が決定してからである。以前に辞書引きし、意味をひいた単語であっても、異なる語義を採用するならばユーザへ説明しなければならない。

また、システムは行動履歴を見て適当なところで文書の読み上げに戻るようにユーザを促す。読み上げ中の文書には直接関係のない語、つまり語義の説明文中の語の意味の質問

を繰り返した場合に、ユーザに文書の読み上げに戻るよう促す。

第5章 関連研究

本章では、関連研究について述べる。

5.1 ダイアログナビ (Dialog Navigator)

大規模な自然言語テキストを知識源として用いる自動質問応答システムに関する研究を行っている。[9] 自然言語テキストを知識源とする自動質問応答システムにおいて、ユーザ質問と知識ベースの間のマッチングは最も重要な基本的技術である。しかし、現在実用化されているシステムのほとんどが採用している。「代表語同士のマッチング」という方法では、自然言語特有の問題である「表現の揺れ」、すなわち同じ事柄を表現するのに多くの言い方があるという問題への対処が困難です。そこで、同義フレーズ辞書などを用いたより柔軟なマッチングの方法について研究を行っている。

自動質問応答システムになされる質問は曖昧なものが多いため、質問と知識ベースをそのままマッチさせるだけでは満足な回答をすることができない。例えば、「エラーが発生した」とような曖昧な質問だと大量のテキストがマッチしてしまい、ユーザがその中から求める答えを探すのは非常に困難である。そこで、曖昧な質問がなされた場合に自動的に聞き返しを行うための方法について研究を行っている。

5.2 会話インタフェースロボットとの状況依存会話

時間や場所のような実世界状況を認識して行動するロボットと人間との会話についての研究。ロボットは、ユーザーを記憶し、過去の情報要求に関する会話の経験に基づいて現在の会話を個人化する。ロボットは顔を持ち、人間との会話にある種の臨場感を与える。また、ロボットは人工感情によって、人間に対して多様な応答を示し、注意や関心を持続させようと試みる。[8]

5.3 ATIS (Air Travel Information System)

アメリカのDARPA (現ARPA) が主導する音声理解プロジェクトにおいて、共通のタスクを用いて複数の研究機関の開発したシステムを競争させている。そのタスクは、飛行機のフライト案内を音声対話によって行うことである。それぞれの機関が開発した異なるシステムが存在するが、基本的にはユーザの音声入力に基づき、フライト情報を格納したデータベースから、必要な情報を検索するシステムである。

5.4 Talkman

人間と自然なインタラクションを行うエージェントを目指して、ソニーコンピュータサイエンス研究所で開発された。人間的な顔を持ち、音声言語で対話するシステムである。Talkman の主なタスクは、ソニーのコンピュータ関連に製品についての応答である。[3]

5.5 飛遊夢

飛遊夢は、NTT コミュニケーション科学基礎研究所で開発した、音声対話インタフェース研究の実験システムである。インターネットからダウンロードした気象情報にもとづいて、天候に関するユーザの質問に答えることができる。[10]

第6章 おわりに

6.1 まとめ

本研究では、文中の語義の説明を行う対話的システムを試作した。認識辞書にあるテンプレートと一致する発話でないと認識できないことや、システムは人間のように常識や柔軟な推論能力を持っていないため、曖昧さが解消されない場合、すぐユーザに質問してくるなど、まだまだ改良点が多い。しかし、必要に応じて情報の追加をしながら問題の解決を目指すことができるという対話の利点を活かしたシステムを作成することができたと思われる。

6.2 今後の課題

6.2.1 認識辞書の動的作成

現段階では、ユーザが意味を知らない語があった場合、どの語の意味が分からなかったのかを直接システムに伝えることはできず、システムが順に提示してくる語にいちいち返答してやらなければならないなど、ユーザは負担を感じてしまう。ユーザが、意味の分からない語を直接伝えられるようにするには、次の事柄の実現が必要になると思われる。

それは読み上げる文書の中から、ユーザに意味を質問される可能性のある語を取り出して、自動的に認識辞書を作成できるようにすることである。また、読み上げる文書が長い場合、ひとつの認識辞書に全ての語を登録することは困難である。そこで、文書を認識辞書に語を登録できる程度の長さに分割し、部分ごとに認識辞書を作成、読み上げる部分に応じて認識辞書を取り替える必要がある。また、語義を説明するときも、ユーザに説明する文に応じて認識辞書を生成することで、語義の説明文中の語も発話でピンポイントで指定できるようになる。

6.2.2 同義語・上位語の判定

システムによる絞り込みでは、他の語から抽出した代表語が語義の説明文中に出現した場合にその語義を採用している。しかし、代表語とまったく同じ語があった場合のみで、同義語や上位語は無視してしまう。代表語と語義の説明文中の語が上位語、または同義語であるかどうかを判定できるならば、システムによって自動的に絞り込むことのできる語が増えると思われる。

6.2.3 文脈の考慮

現段階では辞書引きの対象となる語の付近にある語の語義しか考慮していない。今後は文脈などを考慮した絞り込みにも取り組む予定である。

謝辞

本研究を進めるにあたり、指導教官である長尾確教授には、研究の心構えなど基礎的なことから、ゼミ等を通しての貴重なご意見、論文指導等を賜り大変お世話になりました。

長尾研究室の梶克彦先輩、山根隼人先輩には、論文の書き方、研究の進め方、発表の仕方など多岐にわたる助言を頂きました。さらに、研究環境の充実や行事等、快適な研究室作りをもしていただき、大変お世話になりました。

また、長尾研究室の秘書である兼松英代さんには通常の秘書業務に加え、差し入れや掃除など細やかな心配りをして頂き、深く感謝しています。

また、長尾研究室の学部生の皆様にも、Java やデータベース検索などの様々な質問に答えて頂くなど大変お世話になりました。皆様の協力はとてもありがたいものでした。

最後に、ここまで育ててくれた両親に深く感謝します。

関連図書

- [1] 伊藤 克亘. 対話システム. 自然言語処理-基礎と応用-. pp.281-300, 1999.
- [2] 加藤 恒昭. 音声対話システム. 自然言語処理-基礎と応用-. pp.302-321, 1999.
- [3] 長尾 確. エージェント指向インタフェース. エージェントテクノロジー最前線. pp.63-80, 2000.
- [4] Sun Microsystems, Inc. JAVA(TM) SPEECH API. <http://java.sun.com/products/java-media/speech/>. 1995.
- [5] 奈良先端科学技術大学院大学自然言語処理学講座. 形態素解析システム 茶筌. <http://chasen.aist-nara.ac.jp/index.html.ja>. 1997.
- [6] The Apache XML Project. Apache Xindice. <http://xml.apache.org/xindice/>. 2001.
- [7] W3C. Extensible Markup Language (XML). <http://www.w3c.org/XML/>. 1996.
- [8] NAGAO, Katashi Situated Conversation with a Communicative Interface Robot. The First International Workshop on Intelligent Media Technology for Communicative Reality [Japanese Society for Artificial Intelligence (JSAI)]. 2002.
- [9] 西田・黒橋研究室. ダイアログナビ (Dialog Navigator). <http://www.kc.t.u-tokyo.ac.jp/msnavi/>. 2002.
- [10] 日本電信電話株式会社. 音声対話システム 飛遊夢 (ひゅーむ). <http://www.brl.ntt.co.jp/cs/dug/hyumu/index-j.html>. 2002.