

ユーザ発話文の解析に基づく発話意図認識手法の提案

熊本忠彦 伊藤 昭

郵政省 通信総合研究所 関西先端研究センター

〒 651-24 兵庫県 神戸市 西区 岩岡町岩岡 588-2

E-mail: kuma@karc.crl.go.jp

あらまし 著者らは、自然言語を用いた対話によってユーザ (computer user) の計算機利用を支援する対話型ヘルプシステムの構築を行なっている。ユーザは、通常は計算機上でタスク (task) を進めるが、何らかの障害/問題が生じた時、ヘルプシステムに音声で支援を求めることができる。ヘルプシステムは、ユーザの音声発話を理解し、その時々状況に合わせて適当な応答を生成する。

我々は、ヘルプシステムの代わりに人間アドバイザーが支援するという疑似的な「対話による支援」実験を行ない、ユーザの音声発話を収録した。本論文では、その音声発話文の解析結果に基づいて、発話意図の記述形式を定義し、発話文から発話意図を認識する手法 (発話意図記述を生成する手法) を提案する。

和文キーワード 対話型ヘルプシステム 発話意図認識 話し言葉

A Proposal of Communicative Intention Recognition Method Based on Pattern Analysis of Spoken Japanese Sentences

Tadahiko Kumamoto Akira Ito

Communications Research Laboratory, MPT

588-2 Iwaoka-cho-Iwaoka, Nishi-ku, Kobe 651-24, Hyogo, Japan

E-mail: kuma@karc.crl.go.jp

Abstract The authors have been developing a dialogue-based consultant system that assists a novice computer user in performing tasks on a computer. If support in performing a task is needed, the user can ask the system for help verbally through natural language dialogue. The system analyzes the user utterance and responds appropriately.

We collected utterance data from a dialogue experiment in which a human advisor, instead of the system under development, provided support to a novice computer user (subject). We propose a method of recognizing communicative intentions (CIs) from utterances based on the result of analysis of the utterance data. Our target language is spoken Japanese.

英文 key words Dialogue-based Consultant System, Communicative Intention Recognition, Spoken Japanese Sentences

1. まえがき

著者らは、自然言語を用いた対話によってユーザ (computer user) の計算機利用を支援する対話型ヘルプシステムの構築を行なっている [1][2]。ユーザは、通常は計算機上でタスク (task) を進めるが、何らかの障害/問題が生じた時、ヘルプシステムに音声で支援を求めることができる。ヘルプシステムは、ユーザの音声発話を理解し、その時々状況に合わせて適当な応答を生成する。

ヘルプシステムにユーザ発話を理解させるためには、音声処理技術だけでなく、話し言葉の特性に対応した言語処理技術も要求される [3][4]。伝らは、日常的な話し言葉 (日常会話文) を対象とする解析手法 [4] を提案しているが、汎用性を持たせようとしているため実用性には乏しい。

一方、Guindon は、支援を要請するための発話 (以下、支援要請発話と呼ぶ) には単純かつ限定された言葉 (英語) が用いられることを示している [5]。Guindon の研究は、英語を対象としたものであったが、日本語による支援要請発話にも何らかの制約があると考えられる。

我々は、ヘルプシステムの代わりに人間アドバイザーが支援するという疑似的な「対話による支援」実験を行ない、ユーザの音声発話を収録し、ユーザ発話データベース (書き起こしテキスト) を作成した。また、その書き起こし文を解析し、支援要請発話にはどのような発話意図が現れるのか、発話意図と発話形態の間にはどのような制約があるのか、ということ調べた。本論文では、その解析結果に基づいて、発話意図の記述形式を定義し、発話文から発話意図を認識する手法 (発話意図記述を生成する手法) を提案する。

我々が提案する発話意図記述は、発話意図タイプと発話命題から構成される。発話命題は、タスク領域上で定義可能な「もの」や「行為」、「状態」を表現し、フレーム表現で記述される。発話意図タイプは、ユーザの発話命題に対する態度 (attitude) を表現し、実際の支援要請発話文の解析結果に基づいて、12 タイプに分類される。発話意図認識では、この発話意図タイプを正しく推定することが重要となる。したがって、本論文では、発話意図タイプの推定法を中心に議論する。

今回の対話実験でアドバイザーによってサポートされるプログラムは、UNIX 計算機上の X ウィンドウで動作する電子メール管理プログラム XMH [6] である。XMH は、電子メールを受け取る/表示する/作成する/送る/保管するといった機能を有し、主にマウスで操作される。ユーザは、この XMH を用いて与えられた課題 (お茶会のアレンジなど) を遂行する。

以下、2. で我々が作成したユーザ発話データベースについて述べる。このデータベースは 16 個のデータセット (475 文) を有するが、うち 10 セット (298 文) を解析の対象とし、

残り 6 セット (177 文) を提案する認識手法の性能評価のために用いる。3. および 4. でユーザ発話データベース (10 セット) の解析を行ない、発話意図の記述形式の定義および発話文から発話意図記述を生成する手法の提案を行なう。5. で評価用データセット (6 セット) を用いて発話意図認識手法の性能を評価する。6. で本手法の問題点などについて考察する。7. で本論文の結論と今後の課題について述べる。

2. ユーザ発話データベースの作成

本章では、「対話による支援」実験の様子を述べ、作成された発話データベースの諸元を示す。

「対話による支援」実験は、図 1 のような環境下で行なわれた。被験者は、普段は通常の入力方式 (主にマウス操作によるコマンド入力) でタスクを進め、何らかの困難が生じた時、マイクに向かって発話し、支援を求めた。アドバイザーは、被験者とは離れた別室にいたが、XMH の画面状態やユーザの操作履歴をモニターで見ることができ、その時々状況に応じて適切な応答文を生成した。応答文は、XMH に付加された応答文出力用のヘルプウィンドウに表示される。音声出力ではなく、テキスト出力であるのは、支援内容の記録性を考慮したためである。

ここで、対話実験の仕様を示す。

被験者 : 計算機利用歴のない大学生 16 人
タスク領域 : XMH の操作
平均実験時間 : 1 時間 11 分/人

発話データベースは、次のようにして作成された。まずはじめに、ユーザ発話が収録された録音テープからユーザ発話文を書き起こした。このとき、書き起こしミス (誤字、脱字など) がなくなるまで、繰り返しチェックした。ユーザ発話には、日常会話文と同様、間投表現や言い直し、言い淀み、言い換え表現が現れていたが、取り除いた。このよ

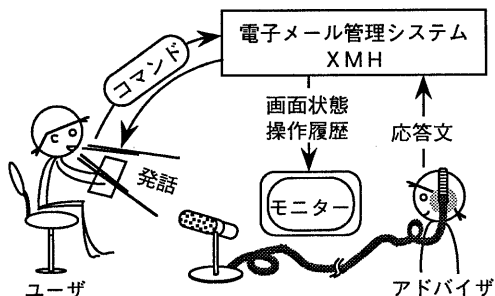


図1 「対話による支援」実験の環境モデル

うな表現の取り扱いは今後の課題である。なお、助詞落ちや助詞の誤運用、倒置などはそのまま残されている。

作成された発話データベースの諸元を示す。

総 発 話 数 : 499 文
総 文 字 数 : 8,147 字 (16.3 字/文)
データセット数 : 16 セット
(被験者 1 名に 1 データセットが対応する)

発話データベースの文構成を示す。

単 文 : 232 文
複 文 : 243 文
フラグメント文 : 24 文
合 計 : 499 文

フラグメント文が約 5% の割合で出現しているが、フラグメント文の処理は基本的に文脈に依存するので、特別な処理が必要とされる [7]。本論文で提案する手法は単文と複文を対象としている。

フラグメント文を除く発話データベース 16 セット (475 文) を解析用データセットと評価用データセットの 2 群に分割した。本論文で解析の対象となるのは解析用データセット (10 セット, 298 文) であり、評価用データセット (6 セット, 177 文) は認識手法の性能評価のために用いられる。各データセット群の構成を表 1 に示す。

ここで、ユーザ発話文のいくつかの例を以下に示す。

「取り込みをして次何番のやつ見たらいいんですか」
「住所忘れました」
「メッセージ作成を押したらいいんですか」
「これはまた入れ直しをしなければいけないんですか」
「メッセージはこれだけでいいんですか」
「これで移動できてるんですか」
「移すのはどうするんですか」
「移動できてないみたいなんですけど」
「スクロールって何ですか」
「この一番上に出たのを消してしまいたいなんですけど」

3. 発話データベースの解析と発話意図記述の定義

本章では、解析用データセット (10 セット, 298 文) の解析を行ない、発話意図の記述形式を定義する。

一般に発話文は、命題部分とモダリティ部分から成ると言われている [8]。我々は、命題部分の意味を表現する記述を“発話命題”と呼び、フレーム表現で記述する [9][10]。また、モダリティ部分には“発話意図を表現するもの”の他にもニュアンスやユーザの感情などが含まれているが、これらをノイズとみなし、無視する。残った“発話意図を表現す

るもの”は、発話命題に対するユーザの態度 (attitude) と考えられる。我々は、このユーザの態度の種類を“発話意図タイプ”と呼び、シンボルで記述する。したがって、一つの発話意図は、

(発話意図タイプ 発話命題)

という形で記述される。

本章の構成は、次のようになっている。3.1 で発話データベースの解析を行ない、支援要請発話にはどのような種類の発話意図が現れるのかを調べ、それぞれの種類に対応する発話意図タイプを定義する。3.2 で発話命題のフレーム表現による記述形式について述べる。

3.1 発話意図タイプの分類

発話意図タイプは、発話意図の種類に対応しており、解析用発話文 298 文の解析結果によれば、表 2 に示す 8 種類に分類される。支援要請発話文における発話意図の種類とは、ユーザがアドバイザーに要求している応答の種類に他ならない。このことは、ヘルプシステムの応答パターンが発話意図タイプの値に応じて選ばれることを意味する。この「応答パターン選択のためのキー」という観点を考慮すると、発話意図タイプ「属性値」と「真偽値」はさらに細分類できる。例えば、発話「メッセージ作成を押したらいいんですか」のようにこれから行なう行為 (未来行為) の正当性を問うような発話や、発話「メッセージはこれだけでいいんですか」のように現在の状態の正当性を問うような発話、発話「これで移動できてるんですか」のように過去の事象に関して問うような発話では、それぞれ異なる知識を用いて応答が生成される。したがって、発話意図タイプ「属性値」と「真偽値」は表 3 のように細分類される。なお、表 2 と表 3 で示されている発話文数は、著者らが以上の観点で解析用発話文 298 文を分類した結果である。

発話意図タイプの記述形式を拡張バックス記法でまとめる。

〈発話意図タイプ〉 ::= 手法 | 〈属性値〉 | 〈真偽値〉 |
概念 | 事実 | 意図 | 対話開始 |
対話終了
〈属性値〉 ::= 属性値 | 属性値:OK
〈真偽値〉 ::= 真偽値 | 真偽値:OK |
真偽値:EQ | 真偽値:NO

3.2 発話命題のフレーム表現による記述

発話命題は、タスク領域上で定義可能な「もの」や「行為」、「状態」を表現し、フレーム表現で記述される。

フレームには、タスク領域内の「もの」を表現する“オブジェクトフレーム”とタスク領域内で可能な「操作」を表現する“イベントフレーム”がある。イベントフレームには、「行為」を記述する“アクションフレーム”と「状態」を記述する“ステートフレーム”がある。各種フレームの例を図 2 に示す。

表1 発話データベースの構成

	番号	文数		番号	文数
解析用データセット	1	20 文	評価用データセット	11	73 文
	2	64 文		12	18 文
	3	59 文		13	37 文
	4	33 文		14	17 文
	5	6 文		15	11 文
	6	46 文		16	21 文
	7	9 文			
	8	20 文			
	9	7 文			
	10	34 文			
	計	298 文		計	177 文

表2 発話意図の種類と発話意図タイプの対応関係

発話意図タイプ	発話意図の種類	発話文数
手法	手法に関する質問	82 文
属性値	属性値に関する質問	27 文
真偽値	真偽値に関する質問	128 文
概念	概念に関する質問	4 文
事実	事実等に関する信念の表明	20 文
ゴール	タスクゴールの表明	4 文
対話開始	対話開始の合図	28 文
対話終了	了解の合図	5 文

表3 発話意図タイプ「属性値」と「真偽値」の細分類

質問内容	真偽値	属性値	サブタイプ
未来の行為/状態の正当性	52 文	8 文	OK
現在の状態の正当性	19 文	0 文	EQ
未来の行為の禁止性	3 文	0 文	NO
過去の行為/状態	54 文	19 文	なし

1つのフレームは、フレーム名やフレーム番号、そしてスロットとスロット値の集合から構成されている。フレーム名は、そのフレームが表している「もの」または「操作」のクラスを示している。フレーム番号は、同じクラスに属する別個の「もの」や「操作」を区別するためのID番号である。そして、スロット・スロット値の対の集合は、それぞれの「もの」または「操作」の属性を表現している。

例えば、「メッセージ作成を押す」という行為は、図2(a)のようなアクションフレームで表現される。このとき、「メッセージ作成」という名称の命令ボタンは、図2(b)のオブジェクトフレームで表現され、アクションフレーム#<命令ボタンを押す #X139EG76>の対象スロットに束縛されている。

発話命題の記述形式を拡張バックカス記法でまとめる。

<発話命題> ::= <命題>
 <命題> ::= フレーム | <and 形式> | <or 形式>
 <and 形式> ::= (<and> <命題> <命題>)*
 <or 形式> ::= (<or> <命題> <命題>)*

4. 発話データベースの解析と発話意図認識手法の提案

本章では、解析用データベース(10セット、298文)の解析に基づいて、発話意図認識手法を提案する。

発話意図認識は、図3に示されるアルゴリズムに沿って処理される。以下、各処理部について述べる。

形態素解析部JUMAN

形態素解析部は、入力発話文の形態素解析を行なう。形態素解析部には京都大学長尾研究室から提供されている形態素解析システムJUMAN[11]を用いている。JUMANの入出力例を表4に示す。JUMANは、入力発話文をいくつかの形態素に分解し、各形態素に原形や品詞名、品詞細分類名、活用形といった情報を付加する。なお、本論文の文法用語は、JUMANで採用されている益岡・田窪文法[12]に依っている。

特徴抽出部

特徴抽出部では、JUMANの出力である形態素列から発話意図タイプ推定用の特徴量と発話命題生成のための情報が抽出される。特徴抽出部の入出力例を表5に示す。

発話意図タイプ推定用特徴量には、「機能語」や品詞情報(品詞名または品詞細分類名)、活用形情報、原形情報の4種類がある。機能語の一覧を表6に示す。また、名詞や指示詞、判定詞からは品詞情報(サ変名詞、普通名詞、形式名詞、固有名詞、副詞の名詞、名詞形態指示詞、判定詞)が抽出される。機能語に含まれていない動詞からは、活用形情報が抽出され、形容詞は状態述語として扱われる。提題助詞や格助詞の「で」、取り立て助詞の「も」、「でも」からは原形情報(とは、って、は、で、も、でも)が抽出される。また、動詞「する」や「できる」、動詞性接尾辞「れる」や「いる」、「る」、「せる」、「ます」、形容詞性述語接尾辞「ない」は直前に抽出された特徴量を別の特徴量に変換する働きがある。変換規則を表7に示す。

発話命題生成のための情報には、原形情報と品詞情報が

#<命令ボタンを押す #X139EG76> —
 対象スロット → #<命令ボタン #X139EB8E>

(a) アクションフレームの例

#<命令ボタン #X139EB8E> —
 名前スロット → メッセージ作成

(b) オブジェクトフレームの例

図2 フレーム表現の例

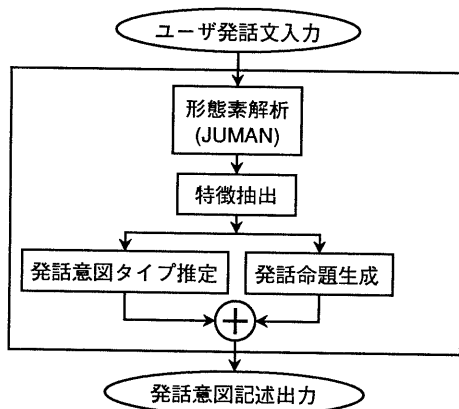


図3 発話意図認識の流れ

あり、名詞や動詞、形容詞、格助詞から抽出される。

発話意図タイプ推定部

発話意図タイプの推定は、入力発話文から抽出された発話意図タイプ推定用特微量と表8に示されるパターン-発話意図タイプ変換テーブルとのパターンマッチングによって行なわれる。

パターン-発話意図タイプ変換テーブルはルールの集合であり、各行が1つのルールに対応している。この変換テーブルは、解析用データセット(10セット、298文)の各文から得られた発話意図タイプ推定用特微量を発話意図タイプ毎に整理した結果、得られた。

パターン-発話意図タイプ変換テーブルの記述形式を拡張バックス記法でまとめる。

```

<ルール> ::= (<パターン部>
               [<発話意図タイプ> | 再帰処理1 |
               再帰処理2 | 再帰処理3])
<パターン部> ::= (<パターン> [& | <パターン>]*)
<パターン> ::= (<特微量>*)
<特微量> ::= <機能語> | <品詞情報> |
              <活用形情報> | <原形情報>
  
```

パターンマッチングは、次のようにして行なわれる。

【ルール r ($1 \leq r \leq R$) の適用】ルール1(変換テーブルの最上行)からルールR(変換テーブルの最下行)まで、順次ルールを適用し、そのパターン部と入力特微量とのマッチングに成功したら、対応する発話意図タイプを発火(fire)する。もしくは、指示された再帰処理(後述)を行なう。

【パターン部のマッチング】パターンのマッチングは、文尾のパターンから文頭方向へと行なわれる。入力特微量のある部分列とパターンが完全に一致すれば、そのパターンのマッチングは成功である。成功したら、その部分列よりも

文頭方向にある特微量を新たに入力特微量とみなし、次のパターンのマッチングを行なう。一番文頭にあるパターンのマッチングに成功したら、そのパターン部のマッチングは成功である。&が現れた場合は、&に続くパターンに対し、入力特微量を初期値(パターン部とのマッチングを開始した時の値)に戻して、パターンマッチングを行なう。また、パターンが空リストの場合は、常にマッチングに成功する。

【再帰処理1】マッチングに成功した特微量よりも文頭方向にある特微量を新たに入力特微量とみなし、ルールの適用をルール1からやり直す。

【再帰処理2】マッチングに成功した特微量よりも文尾方向にある特微量を新たに入力特微量とみなし、ルールの適用をルール1からやり直す。

【再帰処理3】マッチングに成功した特微量を入力特微量から取り除いた後、ルールの適用をルール1からやり直す。

発話命題生成部

発話命題生成部は、発話命題生成のための情報からフレームを生成し、必要に応じてフレーム同士の束縛を行なう。

例えば、発話文「メッセージ作成を押したらいいんですか」から抽出された発話命題生成のための情報「メッセージ作成(固有名詞)」「および」「を(格助詞)」「押す(動詞)」に対して、固有名詞「メッセージ作成」からは、オブジェクトフレーム「命令ボタン」(図2(b)参照)を生成する。ただし、そのフレーム「命令ボタン」の名前スロットには、「メッセージ作成」という値が束縛されている。また、動詞「押す」からは、アクションフレーム「押す」を生成する。フレーム「押す」は、その対象スロットにフレーム「命令ボタン」を束縛する。その結果、フレーム「押す」は、そのサブクラスであるフレーム「命令ボタンを押す」(図2(a)参照)に特化(specialize)される。この例では、格情報「を(格助詞)」は使われなかったが、フレーム同士の束縛関係に曖昧性があるような場合には、格情報が有効に利用される。また、曖昧性がある場合には、格情報が正確に運用されることも支援要請発話文の解析によって示されている[13]。

上述の例文の場合、フレーム「命令ボタンを押す」が発話命題として出力される。

発話意図記述の生成

発話意図タイプ推定部の出力と発話命題生成部の出力とが合成されて、1つの発話意図記述を生成する。

例えば、発話「メッセージ作成を押したらいいんですか」からは、発話意図タイプ推定用特微量として、(文頭 固有名詞 タ系条件形 OK述語 疑問助詞 文末)が抽出され、パターン-発話意図タイプ変換テーブルにより、発話意図タイプ「真偽値:OK」が選ばれる。また、発話命題生成のための情報としては、「メッセージ作成(固有名詞)」,「を(格助詞)」,「押す(動詞)」が抽出され、アクションフレーム#(命

表4 形態素解析システム JUMAN の入出力例

入力発話文「メッセージ作成を押したらいいんですか」
出力形態素列：

形態素	原形	品詞	細分類	活用形
メッセージ	メッセージ	名詞	固有	*
作成	作成	名詞	格助詞	*
を	を	助詞	格助詞	*
押したら	押す	動詞	*	タ系条件形
いい	いい	形容詞	*	基本形
んです	んだ	助動詞	*	デス列
				基本形
か	か	助詞	終助詞	*

表5 特徴抽出の例

発話意図タイプ 推定用特徴量	対応する 入力形態素	発話命題生成 のための情報
文頭←		
固有名詞←	メッセージ作成	→メッセージ作成 (固有名詞)
	を	→を(格助詞)
タ系条件形←	押したら	→押す(動詞)
OK述語←	いい	
	んです	
疑問助詞←	か	
文末←		

表6 機能語の一覧

機能語の名称	対応する語
状態述語	ある, 違う(状態動詞), ない 読める, 消せる(可能動詞) 来る, 消える, 出る(無意志動詞)
教示述語	教える
欠如述語	忘れる, わからない
知識述語	知る, 覚える, わかる
受益述語	もらう, くれる
次	次, 初め, 最初
HOW句	どうする, どうやる, どう
手法語	仕方, 方
疑問語	誰, 何処, 何時(“何”を除く疑問詞)
何	何
OK述語	いい, よい, よろしい
疑問助詞	か, ね, よね, かな, のか
する	する
何でも	何でも(陳述副詞)
禁止述語	駄目だ, いけない
欲求述語	欲しい, たい(形容詞性述語接尾辞)
対話開始合図	すみません, ああ
対話終了合図	わかりました
文頭	(文頭)
文末	(文末)

表7 接尾辞等による特徴量の変換規則

直前の特徴量	入力形態素	特徴量の変換
“サ変名詞”	“する”	(“する”の活用形)
“サ変名詞”	“できる”	“状態述語”
“知識述語”	“ない”	“欠如述語”
“タ系連用テ形”	“ます”	“状態述語”
(活用形)	“ます”	(“ます”の活用形)
(活用形)	“せる”	(“せる”の活用形)
(活用形)	“ない”	“状態述語”
(活用形)	“いる”	“状態述語”
(活用形)	“れる”	“状態述語”
(活用形)	“る”	“状態述語”

表8 パターン-発話意図タイプ変換テーブル

((対話開始合図))	対話開始)
((対話終了合図))	対話終了)
((サ変名詞 欠如述語))	概念)
((形式名詞 欠如述語))	手法)
((手法語 欠如述語))	手法)
((状態述語 欠如述語))	属性値)
((普通名詞 欠如述語))	属性値)
((欠如述語))	再帰処理1)
((普通名詞 教示述語))	属性値)
((HOW句))	手法)
((疑問語)(タ系条件形 OK述語))	属性値:OK)
((疑問語)(基本形 文末))	属性値:OK)
((疑問語))	属性値)
((何 も))	再帰処理2)
((何 する))	手法)
((何)(何 判定詞))	手法)
((次) & (何 判定詞))	手法)
((とは 何 判定詞))	概念)
((って 何 判定詞))	概念)
((何)(タ系条件形 OK述語))	属性値:OK)
((何)(基本条件形 OK述語))	属性値:OK)
((何)(基本形 疑問助詞))	属性値:OK)
((禁止述語))	真偽値:NO)
((タ系条件形 OK述語))	真偽値:OK)
((タ系連用テ形 OK述語))	真偽値:OK)
((基本条件形 OK述語))	真偽値:OK)
((状態述語 OK述語))	真偽値:OK)
((サ変名詞 で OK述語))	真偽値:OK)
((基本形 疑問助詞))	真偽値:OK)
((基本形 文末))	真偽値:OK)
((名詞形態指示詞 で)(OK述語))	真偽値:EQ)
((普通名詞 で)(OK述語))	真偽値:EQ)
((何でも OK述語))	真偽値:EQ)
((疑問助詞))	真偽値)
((タ系連用テ形 受益述語))	意図)
((欲求述語))	意図)
(())	事実)

令ボタンを押す #X139EG76) が生成される。したがって、発話「メッセージ作成を押したらいいんですか」に対する発話意図記述は、

(真偽値:OK #〈命令ボタンを押す #X139EG76)〉)

となる。ただし、#〈命令ボタンを押す #X139EG76)〉の対象スロットには、#〈命令ボタン #X139EB8E)〉が束縛されている。また、#〈命令ボタン #X139EB8E)〉の名前スロットには、“メッセージ作成”という値が束縛されている。

5. 評価実験

本章では、発話意図認識手法の性能を評価用データセット (6 セット, 177 文) を用いて評価する。

4. で示したパターン-発話意図タイプ変換テーブル (表 8) を用いて、発話意図タイプの推定実験を行なった。実験結果を次に示す。

	正答率 (%)
解析用データセット (298 文)	100.0
評価用データセット (177 文)	89.3

また、解析用データセットの数を 1 セットずつ増やしながら、その時々の評価用データセットに対する正答率を調べた。実験結果を次に示す。

解析用データセット数 (セット)	11	12	13
評価用データセット数 (セット)	5	4	3
正答率 (%)	96.2	98.8	100.0

解析用データセットが 13 セット (406 文) のとき、残り評価用データセット 3 セット (69 文) に対して、正答率が 100.0% になった。このときまでに、パターン-発話意図タイプ変換テーブル (表 8) に新たに追加されたルールは、次の 11 ルールである。

(普通名詞 状態述語 欠如述語))	手法)
(文頭 欠如述語))	手法)
(手法語 教示述語))	手法)
(疑問語 でも))	再帰処理 2)
(疑問語)(基本形 疑問助詞))	属性値:OK)
(何)(も))	再帰処理 3)
(何))	属性値)
(固有名詞 で OK 述語))	真偽値:OK)
(基本連用形 で OK 述語))	真偽値:OK)
(も))	再帰処理 3)
(OK 述語))	真偽値:EQ)

この新たに作成されたパターン-発話意図タイプ変換テーブル (計 48 ルール) を用いて、各ルールの利用頻度を全発話文 475 文に対して調べた。そして、利用頻度の高いルールから累積しながら、各時点における正答率を全発話文 475 文に対して調べた。その実験結果を図 4 に示す。

我々が提案している発話意図認識手法は、文の表層的な知識 (原形情報、品詞情報、活用形情報) と単純な意味情報 (機能語) だけを用いて、発話意図タイプの推定を行なっている。しかしながら、解析用データセットに対し、100% の正答率を得たということは、発話意図タイプを推定するのに十分な特徴量が抽出されていることを示している。また、評価用データセットに対しても、約 90% という高い正答率を得たことは、パターン-発話意図タイプ変換テーブルに汎用性があることを示している。

一方、図 4 は、パターン-発話意図タイプ変換テーブルのルールが汎用性の高いルールと個別対応的なルールとに分けられることを示している。このことは、例えば、汎用性の高い 20 ルールを用いれば、未知データの発話意図タイプ推定に対しても、90% 以上の正答率が期待できることを示している。

6. 考察

本章では、発話意図記述、および発話意図認識手法の特徴について考察する。

我々は、発話意図を発話意図タイプと発話命題の対として記述する。このように発話意図をモダリティ部分と命題部分に分けて記述することは、ユーザの発話内容を制限する。発話内容はタスク領域内の事物に限られ、「知っていることを知っている」のようなリフレクティブ (reflective) な表現を用いることができない。しかしながら、ヘルプシステムの対話目的が「ユーザの計算機利用を支援する」であることを考えれば、このような制限も厳しいものではない。

また、発話意図タイプを 12 種類のシンボルで表現するため、ユーザの感情やニュアンスなど、設定されたタイプ以外の発話をヘルプシステムは理解できない。しかしながら、実験結果に示されるようにユーザは当 12 タイプ以外に属する発話をしなかった。これは目的指向型対話が有する本質

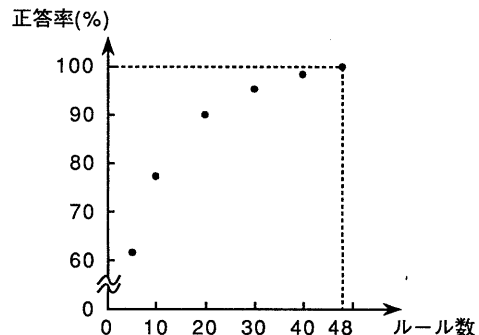


図 4 ルール数と発話意図タイプ推定の正答率

的な性格と考えられる。

久米ら [14] は、発話意図を翻訳に反映できるように「会議問合せ」タスクのための発話行為タイプを7個設定し、さらに、そのタイプ決定をヒューリスティック (社会慣習上の伝達に関する知識を表す規則) や文末表現、副詞類などの大局的な制約を用いて行っている。また、飯田ら [15][16] は、発話の有する対話での機能に着目して約20個の情報伝達行為タイプを設定し、そのタイプ決定を文のモダリティを用いて行っている。彼らの研究は、決定されたタイプの利用に主眼がおかれ、タイプ決定のアルゴリズムは直感的に定義されている。それに対し、我々の手法は実際の発話データの解析結果に基づいて設計されている。

7. まとめ

我々は、ヘルプシステムの代わりに人間アドバイザーが支援するという疑似的な「対話による支援」実験を行ない、ユーザの支援要請発話を収録した。その書き起こしテキストの解析結果に基づいて、本論文では、発話意図の記述形式、および発話意図認識手法 (発話文から発話意図記述を生成する手法) を提案している。

本手法は、発話文中の“支援要請という目的には不要な情報”を無視し、発話命題 (タスク領域中の「行為」や「もの」、「状態」を記述するフレーム) と発話意図タイプ (発話命題に対するユーザの態度 (attitude) の種類) を認識する。このとき、発話意図タイプ推定用の特徴量と発話命題生成のための情報を個別に抽出し、それぞれのアルゴリズムで推定および生成を行なうため、支援要請発話 (話し言葉) に現れる非文法性や不自然さに強い。

本手法を用いて、支援要請発話文の発話意図タイプを推定したところ、学習データ (解析用データセット10セット、298文) に対して100.0%、未知データ (評価用データセット6セット、177文) に対して89.3%の正答率を得た。

現在、新たに支援要請発話文25人分を収集し、解析を行っている。その解析結果を用いて、より汎用的なパターン-発話意図タイプ変換テーブルの作成を目指す。

謝辞

本研究を進める上で、形態素解析システムJUMANは重要な役割を果たしている。話し言葉を形態素解析できるだけの能力を持つJUMANを提供していただいたことに対し、京都大学の長尾真先生、松本裕治先生 (現奈良先端科学技術大学院大学)、およびJUMAN開発グループの方々に深く感謝する。

参考文献

- [1] 伊藤, 樋口, 海老名, 熊本: 対話による計算機利用支援 (1) システムの構成と対話世界モデル, 情報処理学会第42回全国大会, 4C-7, pp. 3-72 - 3-73 (1991).
- [2] A. Ito, T. Ebina and T. Kumamoto: Dialogue-Based User Support for a Visual Interface Software — A Case Study for XMH —, Proc. of Pacific Rim International Conference on Artificial Intelligence, Seoul, Korea (Sep. 1992).
- [3] 飯田 仁: 自然言語対話の言語運用特性と対話処理の研究課題, 人工知能学会誌, 3, 4, pp. 49-56 (July 1988).
- [4] 伝, 飯田: 情報伝達の観点から見た日常会話文の解析手法, 人工知能学会研究会資料, SIG-F/H/K-9101-12(12/6), pp. 111-120 (1991).
- [5] Raymonde Guindon: How to Interface to Advisory System? — Users Request Help with A Very Simple Language —, Proceedings of the CHI'88 Conference, pp. 191-196, (1988).
- [6] J. Peek: MH and XMH — E-mail for Users and Programmers —, O'Reilly & Associates, Inc. (1991).
- [7] J. F. Allen and C. R. Perrault: Analyzing Intention in Utterances, Artificial Intelligence, 15, pp. 143-178 (1980).
- [8] 益岡隆志: モダリティの文法, くろしお出版, 東京 (1991).
- [9] 伊藤, 樋口, 海老名, 熊本: 時間概念を持つフレーム表現の試み, 日本ソフトウェア科学会第8回大会, B3-2, pp. 237-240 (1991).
- [10] A. Ito, H. Higuchi, T. Kumamoto and T. Ebina: A Knowledge Representation for a Dialogue System Understanding Natural Language and Command Input in the Unified Context, Proc. of Japan-Australia Joint Symposium of Natural Language Processing, pp. 95-105, Iizuka, Japan (Oct. 1991).
- [11] 妙木, 松本, 長尾: 汎用日本語辞書および形態素解析システム, 情報処理学会第42回全国大会 (1991).
- [12] 益岡, 田窪: 基礎日本語文法, くろしお出版, 東京 (1991).
- [13] 熊本, 伊藤, 海老名: 対話型計算機利用支援システムにおける音声理解について, 電気関係学会関西支部連合大会, G8-10, p. G268 (1992).
- [14] 久米, 小暮: 発話意図の翻訳のための発話行為推論部, 情報処理学会第39回全国大会, 3F-5, pp. 612-613 (1989).
- [15] 山岡, 飯田: 文脈を考慮した音声認識結果絞り込み手法, 情報処理学会自然言語処理研究会資料, 78-16, pp. 121-128 (1990).
- [16] 飯田, 有田: 4階層プラン認識モデルを使った対話の理解, 情報処理学会論文誌, 31, 6, pp. 810-821 (June 1990).