

発話意図タイプ決定ルールの事例からの学習

熊本 忠彦 伊藤 昭

郵政省 通信総合研究所 関西支所

〒 651-24 兵庫県神戸市西区岩岡町岩岡 588-2
Tel.: (078) 969 2173, E-mail: kuma@crl.go.jp

あらまし 我々は、自由かつ自然な音声対話が可能なシステムの構築を目指している。そのための第一歩として、書き起こされた話し言葉（漢字仮名混じりのテキスト文）から発話意図記述と呼ばれる意味表現を生成することを試みている。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ている。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) の種類を表現し、12タイプに分類される。発話理解では、この発話意図タイプを正しく推定することが重要である。本稿では、事例（発話意図タイプ付ユーザ発話文）から発話意図タイプ決定用のルールを自動抽出する手法をいくつか示し、比較検討する。

キーワード 音声対話システム、支援対話、発話理解、Nグラム

Automatic Extraction of CI Type Determination Rules from Cases

Tadahiko KUMAMOTO and Akira ITO

Kansai Advanced Research Center, Communications Research Laboratory, MPT

588-2 Iwaoka-cho Iwaoka, Nishi-ku, Kobe, 651-24 Hyogo, Japan
Tel.: (078) 969 2173, E-mail: kuma@crl.go.jp

Abstract We are aiming to develop a dialogue system that understands and generates spontaneous speech. A method has already been developed to construct a semantic structure, called a communicative intention (CI) description, from a spoken sentence (kanji-kana mixed sentence transcribed from an utterance). A CI description consists of a CI type and an utterance proposition. An utterance proposition represents “objects,” “actions,” or “states” that can be defined in a task domain, and is described in frame representation. A CI type represents the kind of user’s attitude to the utterance proposition, and there are 12 CI types. In this article, we propose methods for extracting CI type determination rules from cases (user spoken sentences with CI types).

key words Dialogue system, advisory dialogue, spoken-language understanding, N-gram

1. まえがき

我々は、音声対話を通してユーザ (computer user) の計算機利用を支援するシステム (以下、支援システムと呼ぶ) 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/問題が生じたとき、支援システムに助けを求めることができる。支援システムは、ユーザの発話を理解し、その時の対話の状況や計算機の状態に応じて適切な応答を生成する。

このような音声対話システムがユーザフレンドリー (user friendly) であるためには、まず第一にユーザの発話の意味を正しく理解できなければならない。但し、本稿では、「発話理解」とは発話文 (漢字仮名混じりのテキスト文) から発話意図記述 (意味表現) を生成することを意味する。

発話意図記述 (Communicative Intention (CI) description) は発話意図タイプ (CI type) と発話命題 (utterance proposition) から構成されている^[2]。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) の種類を表現し、支援対話における支援要請の種類という観点から 12 タイプに分類されている。

文献 [2] では、ユーザ発話文の発話意図タイプをルールを用いて決定する手法を提案した。このルールは、いくつかのヒューリスティックを導入して、人為的に設計されたものであり、客観性に乏しかった。そこで本稿では、事例 (発話意図タイプ付ユーザ発話文) から発話意図タイプ決定のためのルールを自動抽出する手法を提案する。

なお、支援システムがサポートするタスク領域として UNIX 計算機上の X ウィンドウ下で動作する電子メール処理プログラム XMH^[3] が設定されている。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。

以下、本稿の構成について述べる。2. では、訓練用データや評価用データとして利用される事例 (発話意図タイプ付ユーザ発話文) について述べる。3. では、発話文から抽出される発話意図タイプ決定用の特徴量を定義し、抽出された特徴量の列 (以下、特徴ベクトルと呼ぶ) から発話意図タイプを決定するための手法を示す。4. では、事例から発話意図タイプ決定ルールを抽出するためのアルゴリズムを 3 種類示す。5. では、各アルゴリズムの性能を true error rate (評価用事例に対する error rate) を基準に評価する。最後に 6. では、本稿のまとめと今後の課題について述べる。

2. 事例 (発話意図タイプ付発話文) の収集

我々は、支援システムの代わりに人間コンサルタントが支援するという「対話による支援」実験を行ない、ユーザの自由かつ自然な発話を収集した。そして、収集されたユーザ発話を書き起こすことによって、支援対話データベース (ユーザ 42 人文の発話文 855 文を含むテキストデータベース) を作成した^[4]。

この支援対話データベースからユーザ発話文 (フラグメント文 24 文を除くユーザ 16 人分の発話文 475 文) を抽出し、ユーザ発話文コーパスを作成した。ユーザ発話文には、日常会話文と同様、間投表現や言い直し、言い淀み、言い換え表現などが現れていたが、これらはすべてコーパス作成にあたって取り除かれた。このよ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、倒置なども観測されたが、これらはそのまま残されている。

以上のようにして収集されたユーザ発話文に発話意図タイプを付与した。発話意図タイプには、ユーザ発話に対する応答生成パターン (発話命題の処理手順) を選択するためのマーカー (marker) としての役割が与えられており、支援要請の種類と表 1 に示されたような対応関係にある。未来事象の正当性に関する質問とは「～したらいいんですか」のようなある事象の容認可能性を問うことによって許可を求める文のことであり、未来事象の禁止性に関する質問とは、「～してはいけませんか」のような禁止の形式で問うことによって許可を求める文のことである。また、事象そのものに関する質問とは XMH 上で生じた事象 (出来事/状態) に関する質問のこと

表 1 発話意図タイプと支援要請の種類

| 発話意図タイプ | 支援要請の種類 |
|---------|--------------------|
| 手法 | 手法に関する質問 |
| 概念 | 概念に関する質問 |
| 属性値 | 属性値に関する質問 |
| 属性値:OK | 事象そのもの 未来事象の正当性 |
| 真偽値 | 真偽値に関する質問 |
| 真偽値:OK | 事象そのもの 未来事象の正当性 |
| 真偽値:NO | 未来事象の禁止性 |
| 真偽値:EQ | その他 |
| 信念 | 事実等の表明 |
| ゴール | タスクゴールの表明 |
| 対話開始 | 対話開始の合図 |
| 対話終了 | 対話終了の合図 |

であり、その他のことに関する質問とは XMH 上のある状態の正当性に関する質問のことである。ここで、各発話意図タイプに対応するユーザ発話文の例を、ユーザ発話文コーパスから抜粋し、表 2 に示す。

3. 発話意図タイプ決定法

3.1 発話意図タイプ決定用特微量の定義

発話意図タイプ決定用特微量には、「意味語 (semantic word)」や品詞情報 (品詞名/品詞細分類名)、活用形情報、原形情報の 4 種類がある。意味語は、本稿では発話意図タイプの決定に影響を与える意味情報の種類と定義され、入力形態素の原形情報と品詞情報から辞書引きされる。例えば、話者の「良い」という判断を表す形容詞「いい、よい、…」からは意味語「OK 述語」が抽出され、疑問を表す終助詞「か、ね、…」からは意味語「疑問助詞」が抽出される。ここで、意味語の一覧を表 3 に示す。

一方、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、副詞の名詞、名詞形態指示詞、判定詞) が抽出され、動態動詞からは活用形情報が抽出される。また、提題助詞 (「とは」、「って」、「は」など) 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。

3.2 発話意図タイプ決定法

形態素解析 & 特徴抽出:

入力発話文は、まず汎用日本語形態素解析システム JUMAN^[5]によって形態素解析される。JUMAN は、入力発話文を適当な形態素に分解し、各形態素に原形情報や品詞情報、活用形情報などを付加する。なお、本稿の文法用語は、JUMAN で採用されている益岡・田窪文法^[6]に依っている。

発話意図タイプ決定用特微量は、JUMAN の出力である形態素列から抽出される。例えば、発話文「終了を押せばいいんですね」からは特徴ベクトル「(サ変名詞 動態述語 基本条件形 OK 述語 疑問助詞 文末)」が得られる。この様子を表 4 に示す。

発話意図タイプの決定:

入力発話文の発話意図タイプは、その発話文から抽出された特徴ベクトルに発話意図タイプ決定ルールを適用することによって決められる。ここで文献 [2] で得られたルールの一部を、例として表 5 に示す。また、ルールの記述形式を拡張バックス記法で示す。
 〈ルール〉 ::= (〈パターン部〉 〈発話意図タイプ〉)
 〈パターン部〉 ::= (〈パターン〉 〈パターン〉)*
 〈パターン〉 ::= (〈特微量〉 〈特微量〉)*
 〈特微量〉 ::= 〈意味語〉 | 〈品詞情報〉 | 〈活用形情報〉 | 〈原形情報〉

ルールの適用は、次のようにして行なわれる。

表 2 発話意図タイプとユーザ発話文の例

| | |
|--------|-------------------|
| 手法 | 移すのはどうするんですか |
| 概念 | スクロールって何ですか |
| 属性値 | 住所を忘れました |
| 属性値:OK | 何を書いたらいいんですか |
| 真偽値 | これで移動できてるんですか |
| 真偽値:OK | 終了を押せばいいんですね |
| 真偽値:NO | これ押さなくちゃ駄目なんです |
| 真偽値:EQ | メッセージはこれだけでいいんですか |
| 信念 | 移動できてないみたいなんですけど |
| ゴール | ロイの次にカーソルを持っていきたい |
| 対話開始 | すみません |
| 対話終了 | わかりました |

表 3 意味語の一覧

| 意味語 | 対応する語 (代表的な例) |
|--------|---|
| 対話開始合図 | すみません、あの |
| 対話終了合図 | わかりました |
| HOW 句 | どうする、どうやる、どう |
| 教示述語 | 教える |
| 受益述語 | もらう、くれる |
| 欲求述語 | 欲しい |
| 疑問語 | 何処 (「何」を除く疑問詞) |
| 何 | 何 |
| OK 述語 | いい、よい、よろしい |
| 欠如述語 | 忘れる、わからない |
| 禁止述語 | 駄目だ、いけない |
| 疑問助詞 | か、ね、よね (終助詞) |
| 文末 | (文末) |
| 手法語 | 仕方、方 |
| する | する |
| 何でも | 何でも (陳述副詞) |
| 動態述語 | 読む、消す (動態動詞) |
| 状態述語 | ある、違う (状態動詞) おかしい、ない (形容詞) 読める、消せる (可能動詞) 消える、出る (無意志動詞) |
| 知識述語 | 知る、覚える、わかる |

【ルール r ($1 \leq r \leq R$) の適用】 入力特徴ベクトルにルール 1 からルール R を順次適用する。入力特徴ベクトルとルール r のパターン部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) し、ルールの適用を終了する。

【パターン部のマッチング】 パターン部の最左パターンから最右パターンへとマッチングを行ない、全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。

【パターンのマッチング】 当該パターンが入力特徴ベ

表4 発話意図タイプ決定用特徴量の抽出

| 入力形態素 | 特徴量 |
|-------|-------------|
| 終了 | サ変名詞 |
| を | |
| 押せば | 動態述語, 基本条件形 |
| いい | OK 述語 |
| んです | |
| ね | 疑問助詞 |
| | 文末 |

表5 発話意図タイプ決定ルール(一部)^[2]

| | |
|-----------------|---------|
| ((疑問語)(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) |

クトル中のある部分列と同一であるならば, 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末方向にある特徴量の列を新たに入力特徴ベクトルとみなし, 次のパターンのマッチングを行なう。

4. Nグラムを用いた発話意図タイプ決定ルールの自動抽出

本章では, 事例(特徴ベクトルと発話意図タイプとの組)から発話意図タイプ決定ルールを自動抽出する手法について述べる。本手法に対する入力事例の集合であり, 出力はルールの順序付集合となる。アルゴリズム a (全探索法)

(1) Nグラムの作成

入力中の各特徴ベクトルからNグラムを作成する。例えば, 3個の特徴量を有する特徴ベクトル「(動態述語 基本形 文末)」からは以下のようなNグラム統計量(3つの uni-gram, 2つの bi-gram, 1つの tri-gram)が得られる。

- (動態述語) (基本形) (文末)
- (動態述語 基本形) (基本形 文末)
- (動態述語 基本形 文末)

(2) 頻度の計数

各異なりNグラム統計量に対し, 発話意図タイプ毎に頻度を数える。例えば, 入力を3個の事例

- ((状態述語 文末) 信念)
- ((動態述語 タ形 文末) 信念)
- ((動態述語 基本形 文末) 真偽値:OK)

と仮定すると, 各異なりNグラム統計量に対し, 発話意図タイプ毎に以下のような頻度が得られる。

- ((動態述語 基本形 文末) (真偽値:OK . 1))
- ((動態述語 タ形 文末) (信念 . 1))

- ((動態述語 基本形) (真偽値:OK . 1))
- ((動態述語 タ形) (信念 . 1))
- ((状態述語 文末) (信念 . 1))
- ((基本形 文末) (真偽値:OK . 1))
- ((タ形 文末) (信念 . 1))
- ((動態述語) (真偽値:OK . 1) (信念 . 1))
- ((状態述語) (信念 . 1))
- ((基本形) (真偽値:OK . 1))
- ((タ形) (信念 . 1))
- ((文末) (真偽値:OK . 1) (信念 . 2))

(3) 最適Nグラム統計量の抽出

該当する異なりNグラム統計量が1つになるまで, 以下の操作を順に行なう。

3-1) 発話意図タイプ毎に計数された頻度の中の最大頻度が閾値 e 以上のものを選ぶ。もし, 該当するNグラム統計量がなければ, 見つかるまで閾値 e を1ずつ下げる。

3-2) 発話意図タイプの種類数が最少のものを選ぶ。

3-3) 最大頻度が最大のものを選ぶ。

3-4) 特徴量数が少ないものを選ぶ。

3-5) 先に抽出されたものを選ぶ。

例えば, (2) で用いた例において, e = 2 のときは異なりNグラム統計量「(文末)」が選ばれる。

(4) 発話意図タイプ決定ルールの生成

選ばれた異なりNグラム統計量を, 対応する発話意図タイプの中で最大頻度のものと組み合わせ, ルール r として登録する。例えば, 上述の例では, 「(((文末)) 信念)」というルールが得られる。但し, r は正の整数であり, 抽出された順序を表す。

(5) 終了条件のテスト

生成されたルールにマッチングする事例を入力から削除する。入力がなくなれば(6)に進み, そうでなければ(1)に進む。

(6) ルールの洗練

まず, 最後に抽出されたルール R のパターン部を「((文末))」にする。次に, r = R, ..., 2, 1 の順に「ルール r - 1 とルール r の発話意図タイプが同じであり, かつルール r - 1 にマッチングする特徴ベクトルがすべてルール r にもマッチングするとき, ルール r - 1 を削除する」という操作を行なう。この結果が出力となる。

アルゴリズム b (意味語の有無を基準にした部分空間探索法)

(i) 意味語の有無を基準にした部分集合の作成

各意味語毎に, その意味語を含む事例を入力から抽出し, 部分集合を作る。このとき, 1個の事例が複数の部分集合に重複して含まれてもよい。意味語は表3に示したように19種類あるので, 19個の部分集合が作られる。

(ii) 種類数と頻度の計数

各部分集合毎に、観測される発話意図タイプの種類数と発話意図タイプ毎の頻度を求める。

(iii) 優先意味語の選定

該当する部分集合が1つになるまで、以下の操作を順に行なう。

iii-1) 発話意図タイプ毎に計数された頻度の中の最大頻度が閾値 e 以上のものを選ぶ。もし、該当する部分集合がなければ、見つかるまで閾値 e を1ずつ下げる。

iii-2) 発話意図タイプの種類数が最少のものを選ぶ。

iii-3) 最大頻度が最大のものを選ぶ。

iii-4) 事例数が最大のものを選ぶ。

iii-5) 先に得られたものを選ぶ。

以上の結果選ばれた部分集合に対応する意味語が優先意味語となる。

(iv) Nグラムを作成

入力中の各特徴ベクトルからNグラムを作成する。

(v) 優先意味語を含むNグラム統計量の抽出

作成されたNグラムから優先意味語を含むNグラム統計量を抽出する。

(vi) 発話意図タイプ決定ルールの抽出

抽出されたNグラム統計量に対し、アルゴリズム a の手順 (2)~(4) を行ない、ルールを抽出する。

(vii) 終了条件のテスト

生成されたルールにマッチングする事例を入力から削除する。入力中に優先意味語を含む事例があれば (iv) に進み、なければ (i) に進む。入力が無くなったら、ルールの洗練 (アルゴリズム a の手順 (6)) を行ない、実行を終了する。

アルゴリズム c (意味語の有無と発話意図タイプの種類とを基準にした部分空間探索法)

(a) 優先意味語の決定

アルゴリズム b の手順 (i)~(iii) に従い、優先意味語を決定する。

(b) 発話意図タイプの順序付け

入力から優先意味語を含む事例を抽出し、各発話意図タイプの出現頻度を調べる。そして、各発話意図タイプに対し、頻度の小さい順に優先順位を付ける。以下、優先順位1位の発話意図タイプを当該発話意図タイプと呼ぶ。

(c) Nグラムを作成

入力中の各特徴ベクトルからNグラムを作成する。

(d) 優先意味語を含むNグラム統計量の抽出

作成されたNグラムから優先意味語を含むNグラム統計量を抽出する。

(e) 頻度の計数

各異なりNグラム統計量に対し、発話意図タイプ毎に頻度を数える。

(f) 最適Nグラム統計量の抽出

該当する異なりNグラム統計量が1つになるまで、以下の操作を順に行なう。

頻度(発話文数)

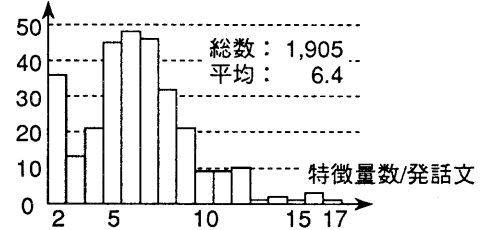


図1 一発話文から抽出される特徴量の数 (訓練用298事例に対して)

頻度(異なりNグラム統計量の数)

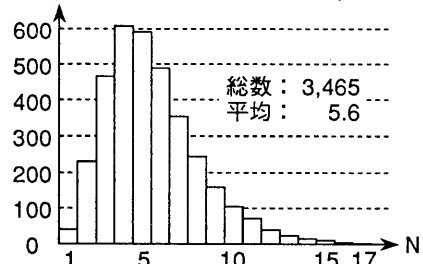


図2 異なりNグラム統計量の分布 (訓練用298事例に対して)

表6 True error rate とルール数

| 手法 | true error rate | ルール数 |
|------------|-----------------|------|
| アルゴリズム a | 0.198 | 63 |
| アルゴリズム b | 0.107 | 68 |
| アルゴリズム c | 0.096 | 44 |
| 文献 [2] の手法 | 0.107 | 29 |

f-1) 当該発話意図タイプの頻度が、その異なりNグラム統計量における最大頻度であり、かつ閾値 e 以上のものを選ぶ。もし、該当するNグラム統計量がなければ、見つかるまで閾値 e を1ずつ下げる。閾値 $e = 1$ まで下げても見つからないときは、当該発話意図タイプを一時的に次の優先順位のものにする。

f-2) 対応する発話意図タイプの種類数が最少のものを選ぶ。

f-3) 最大頻度が最大のものを選ぶ。

f-4) 特徴量数が少ないものを選ぶ。

f-5) 先に抽出されたものを選ぶ。

(g) 発話意図タイプ決定ルールの生成

選ばれた異なりNグラム統計量を、当該発話意図タイプと組み合わせ、ルール r として登録する。

(h) 終了条件のテスト

生成されたルールにマッチングする事例を入力から削除する。入力中に優先意味語を含む事例があれば (c) に進み、なければ (a) に進む。但し、優先意味語を含む事例の中に当該発話意図タイプを含むも

のなければ、次の優先順位をもつ発話意図タイプを当該発話意図タイプとする。入力が無くなったら、ルールの洗練(アルゴリズム a の手順 (6))を行ない、実行を終了する。

5. 提案手法の評価と考察

文献 [2] と同じ条件で評価実験を行なった。すなわち、事例(発話意図タイプ付発話文)を訓練用事例(298 事例)と評価用事例(177 事例)に分け、訓練用事例を入力とし、ルールの抽出(閾値 $e = 1$)を行なった。ここで、入力の性質として、一発話文から抽出される特徴量数の分布を図 1 に、異なり N グラム統計量の N に対する分布を図 2 に示す。また、抽出されたルールの true error rate(評価用事例に対する error rate) とルール数を表 6 に示す。

次に、各アルゴリズムの性能をより客観的に調べるために 5-fold cross-validation に基づく評価実験を行なった。各実験における訓練用事例数は 380、評価用事例数は 95 である。各実験において抽出されたルールの true error rate とルール数の平均値を、閾値 $e = 1, 2, \dots, 10$ においてそれぞれ求めた。結果を図 3 に示す。

アルゴリズム a は、全解探索法であるため、文法的に無意味なルールを抽出しやすい。アルゴリズム b は、意味語を基準に探索空間を限定したため、アルゴリズム a よりも良かったが、その探索空間の中では全解探索であったため、ルールの抽出が非効率であった。アルゴリズム c は、意味語を基準に限定された探索空間の中で、頻度の小さい発話意図タイプに対し、優先的にルールを生成した。これは言い換えると、汎用的なルールを先に抽出することによって、汎用的なルールの抽出を行なったといえる。その結果、より良い性能を得ることができた。

6. まとめ

本稿では、事例(発話意図タイプ付ユーザ発話文)から発話意図タイプ決定ルールを自動抽出するための手法を提案した。1 つのルールは、パターン部と発話意図タイプからなり、パターン部は N グラム統計量(特徴量の列)の形をしている。

今後の課題として、離散型の N グラム統計量を導入してみることに、および抽出されたルールの性質について分析を行なうことの 2 点が挙げられる。

参考文献

[1] 熊本, 伊藤, 海老名: “電子メール利用支援システム”, 通信総研第 87 回研究発表会予稿, 4, pp.

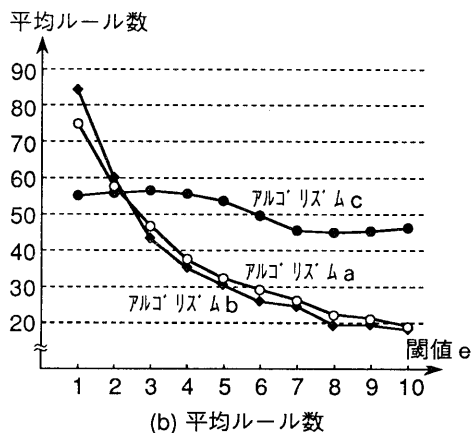
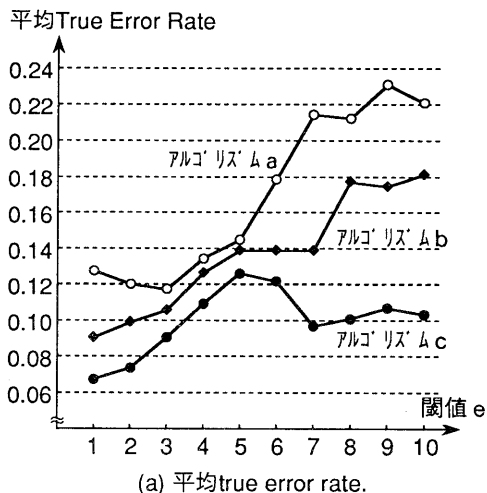


図 3 閾値 e に対する平均 True Error Rate と平均ルール数(5-fold cross-validation; 訓練事例数: 380, 評価事例数: 95)

36-43 (1994).

- [2] 熊本, 伊藤, 海老名: “支援対話におけるユーザ発話意図の認識”, 信学論 D-II, **J77-D-II**, 6, pp. 1114-1123 (1994).
- [3] Peek, J.: “Mh and Xmh - E-mail for Users and Programmers -”, O'Reilly & Associates, Inc. (1991).
- [4] Kumamoto, T., Ito, A. and Ebina, T.: “Design and Construction of an Advisory Dialogue Database,” IEICE Trans. Information and Systems, **E78-D**, 4, pp. 420-427 (1995).
- [5] 妙木裕, 松本裕治, 長尾真: “汎用日本語辞書および形態素解析システム”, 第 42 回情処全大, 1C-9, p. 3-17 (1991).
- [6] 益岡隆志, 田窪行則: “基礎日本語文法”, くろしお出版, 東京 (1991).