

話者の意図理解に基づく対話応答システム Hermes

星野高志[†] 坂本一樹^{††} 原田実^{††}

近年、言葉によってコミュニケーションをする対話応答システムの研究が盛んになっている。しかし、高度なコミュニケーションを行うためには、発話からユーザの意図を理解し、ユーザの要求に応じることが必要になるが、既存の研究ではそれが実現できていない。本研究では、発話文からユーザの意図を理解し、これに応じる行為を実行する対話応答システム Hermes を開発した。Hermes は、話し手の意図を、ユーザ発話から意味解析システム SAGE によって得られるモダリティと語意の組み合わせとして認識する。話し手の意図に対する応答行為としては、対話記録、新聞データ、WEB を知識源とした質問応答や、天気や道案内やスケジュール管理などの行為を実装し、電子秘書として有効なシステムを目指した。

Dialogic answering system Hermes understanding intention of speakers

Hoshino Takashi[†] Sakamoto Kazuki^{††} Minoru Harada^{††}

Recently, the research of the natural language dialogue system is active. However, it cannot be achieved in the proposed systems to understand the user's intention from the utterance and to respond to it. We developed the natural language dialogue system Hermes which understands a user's intention from his/her utterance and do the proper actions corresponding to it. Hermes recognizes the user's intention as a combination of the modality and the meaning of a word which are given by the semantic analysis system SAGE from the user's utterance. Hermes aims at an electronic secretary by performing response acts such as question answering based on the knowledge of the conversation records, the newspaper data, and WEB, weather forecast and traffic guide, scheduling management, etc.

[†] 青山学院大学大学院理工学研究科

^{††} 青山学院大学理工学部情報テクノロジー学科

1. 序論

1.1. 研究背景

近年、言葉によってコミュニケーションをする対話応答システムの研究が盛んになっている。しかし、高度なコミュニケーションを行うためには、発言からユーザの意図を正しく理解し、ユーザの様々な要求に応じることが必要になるが、既存の研究ではそれが実現できていない。MIT のジョセフ・ワイゼンバウムが 1966 年に開発した対話システム ELIZA[1] をきっかけに多くの対話システムが開発されてきた。ELIZA は、ユーザの入力文から取り出した単語を予め用意された文章に組み込むことで質問文を生成し、見かけ上は自然な対話を行うことができる。ELIZA の問題点は、たとえ話し手が自然な対話が成立していると感じていたとしても、ELIZA はユーザの発言の意味を理解しておらず、単語の組み換えを行っているだけという点である。また ELIZA はユーザの発言に対してすべて質問文の形で返すので、話し手の要求に応じるなどといった行為が不可能である。1968 年から 1970 年にかけてテリー・ウィノグラードが SHRDLU[2] を開発した。ユーザは SHRDLU に対し、端末の画面の中の「積み木の世界」に存在するブロック、円錐、球などを動かすよう自然語で命令することができる。SHRDLU の問題点は、「積み木の世界」のような極めて単純な世界以外での応用が困難な点である。単純な世界であるからこそ 50 種類程度の単語で世界の状況を記述することができるが、現実の複雑な世界を記述することは困難である。2008 年に原田らは物語との対話システム Eve[3] を開発した。Eve は対話に必要な知識と推論ルールを自然語のみで記述し、推論エンジンに目標スタックの概念を利用している。ユーザの要求を目標スタックに積み、この目標を推論ルールによって部分目標に展開する。これは部分目標が Eve によって実行可能な原始的な要求、「調査する」や「間違いを正しくする」に展開されるまで繰り返される。そして実行可能な要求をすべて消化し、目標スタックが空になると対話を終了する。Eve の問題点としては、より多くの対話を実現するためには多くの知識や推論ルールを人手で用意する必要があるという点である。現在稼働中のシステムとしては、NTT ドコモの音声エージェント機能のしゃべってコンシェルや、Apple の iOS 向け他機能アプリケーションの Siri が挙げられる。これらのシステムでは、簡単な問い合わせや案内が、スマートフォンアプリケーション上での自然語による対話処理として可能である。このように、対話応答システムは我々にとって身近な存在へととなっている。

1.2. 研究目的

前述の様な状況を考えると、高度な対話応答システムを実現するためには、ユーザの要求に応じる方法と対話に必要な知識の取得において新しい仕組みが必要であると考えた。そこで本研究では、話し手の意図を理解することで話し手が受け手に何を求めているのかを理解し、ユーザを支援する行為を実行することでユーザの役に立つ対話応答システム Hermes を構築する。実行する行為としては計算機上で行える行為に限定し、対話記録を知識源とした質問応答による記憶支援や、新聞データ、WEB を知識源とした質問応答、スケジュール管理、PC の操作支援などを行い、ユーザを支援する。対話記録を知識源とすることによって人手で知識を追加することなくユーザの質問に答えることを目指す。

2. 対話応答システム Hermes のシステム概要

2.1. システム概要

対話応答システム Hermes は、図 1 のように発話文からユーザの意図を理解する意図理解エンジンと、これに応じた行為を実行する行為実行エンジンから成るシステムである。ユーザの発話に含まれる意図と Hermes が実行できる行為を結びつけることによって、ユーザの意図に応じた行為の実行を可能にする。これにより、対話記録や新聞データ、WEB を知識源とした質問応答や、PC 操作、スケジュール管理などのユーザを支援する行為が実行可能である。

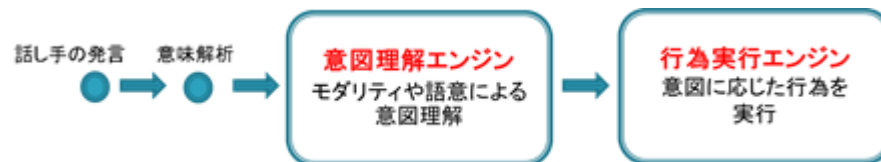


図 1 対話応答システム Hermes の概要

2.2. SAGE による意味解析

Hermes は、原田研究室の意味解析システム SAGE[4]を用いる。SAGE は、日本語を文節の(形態素の)意味や深層格(他の文節との役割関係)などを保持したリストとして表現する。これは、文節を頂点、係り受け関係にある文節間の深層格を辺と考えると、図 2 のような意味グラフとして表現される。語意には EDR 辞書の約 40 万個の概念 ID を用いている。現状の精度は、語意 96%、深層格 93%である。

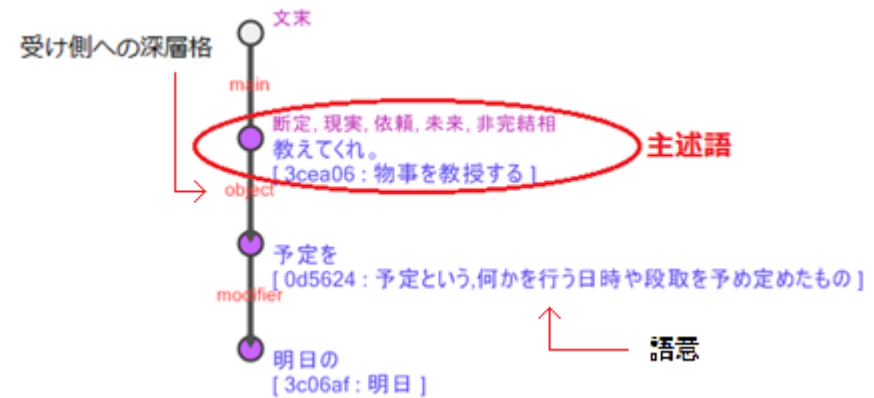


図 2 SAGE の意味解析結果を示す意味グラフ

2.3. Metis による質問応答

Hermes は、原田研究室の質問応答システム Metis[5]を用いる。Metis は SAGE の出力結果である意味グラフに基づいている。意味グラフとは、文節がノード、文節間の深層格が辺に対応した有向グラフである。Metis の質問応答の基本的考え方は、自然言語で入力された質問文と知識文を SAGE で意味グラフに変換しその結果の 2 つのグラフを照合し、両グラフの共通部分グラフの大きさによって 2 つの文の類似性を算出する。そして、質問グラフの質問箇所に対応する知識グラフのノードを回答として抽出し、抽出した回答の正当性をグラフの類似度に基づいて評価する。本研究では、意図「質問」においてユーザに回答を提示するために使用する。質問内容が一般的な内容であれば新聞データまたは WEB 検索結果を知識源とし、個人的な質問であればユーザの対話記録を知識源として回答抽出を行う。NTCIR の Factoid 型質問のテストコレクション QAC1 と CLQA2 の回答精度はそれぞれ 71.9%、77.0%である。また、Non-Factoid 型質問のテストコレクション QAC4 における Metis の精度は 60.0%である。

3. 意図理解エンジン

3.1. 意図とは

本研究における意図とは、話し手が受け手に対して何を求めているかを表すものである。Hermes は、意図を推論することによってユーザの要求を判別する。日本語には豊富な言語表現があるので、話し手の意図が一意に定まる場合でもその意図を表す発言は多くのパターンが考えられる。話し手の意図は発言中の主述語(文末の述語節)のモダリティと語意の組み合わせによって表現されるので、それらを利用することで多様な言語表現の中から話し手の意図を特定することが可能となる。話し手の意図には「受け手にある行為をすることを求める」「受け手に回答を求める」「話し手の意見を受け手に伝達する」などがあり、すべて話し手と受け手との一対一の対話を想定している。話し手の発言の意図と、受け手つまり Hermes が実行出来る行為を結びつけることによって意図に応じた行為の実行を可能とする。本研究では話し手の意図を表 1 のように 9 種類定義した。

表 1 意図の定義

意図	意図の定義
要求	受け手にある行為をすることを求める
質問	受け手に回答を求める
禁止	受け手にある行為を禁止する
表明	話し手の意見を受け手に伝達する
叙述	話し手が観測したことを受け手に伝達する
挨拶	受け手に挨拶する
肯定	受け手による問い返しを肯定する
否定	受け手による問い返しを否定する
その他	相槌など意図を持たないもの

3.2. モダリティと語意について

文は「命題」と呼ばれる客観的な事柄を表す領域と、「モダリティ」と呼ばれる話し手の命題に対する主観的認識や発言態度を表す領域から構成される。モダリティは命題述部の語尾に現れる。

「会議の予定を追加したい」

「会議の予定を追加してほしい」

命題 モダリティ

上の例文「会議の予定を追加したい」「会議の予定を追加してほしい」は異なる文であるが、図 4 で示しているように語意が同一であるので、モダリティを用いなければその差異を表現することが出来ない。SAGE ではモダリティは話し手の命題に対する主観的認識を表す「判断のモダリティ」と発言態度を表す「発言のモダリティ」、命題実現の難易度を表す「程度のモダリティ」という三つのカテゴリに分けており、本研究でもこれを利用する。また、本研究における語意とは、EDR 辞書の概念 ID を指す。

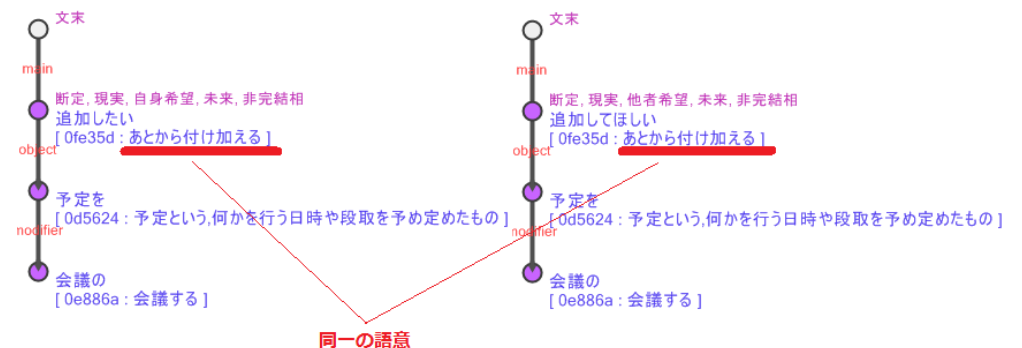


図 4 同一の語意

3.3. モダリティや語意による話し手の意図分類

本研究における話し手の意図は発言中の主述語に含まれるモダリティと語意によって表現される。話し手の発言の主述語にはモダリティが含まれない場合があるので、発言がモダリティを含むか含まないかで意図理解の方法が異なる。主述語にモダリティが含まれる場合には、モダリティによって意図を「要求」「質問」「禁止」「表明」「叙述」のいずれかに分類する。主述語にモダリティが含まれない場合には、語意によって意図を「挨拶」「肯定」「否定」「その他」のいずれかに分類する。また、「質問」と「要求」へ分類する際は、モダリティによる分類を行った後、さらに語意による再分類を行う。これにより、意図の誤分類が減少し、意図理解精度の向上を実現した。

4. 行為実行エンジン

行為実行エンジンでは、意図ごとに Hermes が実行すべき行為を図 5 のように定義した。なお、意図「禁止」「表明」「叙述」「その他」に対して Hermes が実行する行為は現時点では、簡単な返事を返すものであり、今後語意と組み合わせる行為を再定義する必要がある。また、意図「要求」においてはさらに語意によって具体的な行為へ分類し実行する。具体的な行為としては、現時点ではパソコンの操作支援・スケジュール管理・天気予報・道案内機能を実装している。

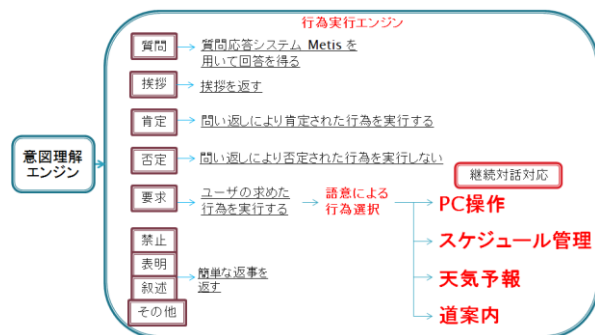


図 5 行為実行エンジン

4.1. 質問

意図「質問」においては、質問応答システム Metis を利用することでユーザーの質問に対して回答を提示する。本来 Metis が利用する知識源は新聞データと WEB のみだが、本研究においてはさらに対話記録を知識源として利用する。この対話記録を知識源とした質問応答によってユーザーの記憶支援を行う。ユーザーからの質問は「個人的な質問」と「一般的な質問」の二つのタイプに大別できる。個人的な質問とはユーザーが過去に Hermes と対話した内容に関する質問のことであり、この質問の場合には対話記録を知識源として回答を抽出する。一般的な質問とは個人的な質問以外の質問であり、この質問の場合には新聞データまたは WEB を知識源として回答を抽出する。一般的な質問において知識源として新聞データと WEB のどちらを利用するかはユーザーが選択する形をとる。

個人的な質問

「自己PRでは何を話そうとしてたっけ？」

「研究の背景はどうでしょうか？」

一般的な質問

「マレーシアの首都と言えば？」

「明石海峡大橋の愛称は何ですか。」

図 6 個人的・一般的な質問

個人的な質問と一般的な質問のどちらの質問なのかによって利用する知識源が異なるのでこれらを判別する必要がある。個人的な質問とは、ユーザーが過去に言及したことに関する質問なので、過去のユーザーの発言をもとに判別を行う。そのため Hermes では対話記録とは別にユーザーの発言の一部を記録している。記録するのはユーザーの発言において、品詞が名詞か動詞である形態素の語意である。質問文中の形態素の語意のうち、語意の記録に含まれている数が閾値以上であれば個人的な質問とし、閾値未満であれば一般的な質問とする。つまり、質問文中に過去にユーザーによって使われた語意が多ければ、個人的な質問として判定される。個人的な質問、一般的な質問のどちらの場合でも、回答を複数取得した場合には、最も総合類似度の高い回答をユーザーに提示する。回答が 1 件も取得できなかった場合には、知識源を変更して質問応答を再度行うかについてユーザーへ問い返しを行う。

ユーザーの発言は、問い返しに対するユーザーの応答を除いて全てデータベースに対話記録として記録する。ここで記録された対話は質問応答システム Metis による質

問応答の知識源として利用する。ユーザの発言は日付 (DateTime)、SAGE 解析結果 (SG)、発言 (Utterance)、意図分類の結果 (Classification) の四つのカラムからなるテーブル、DialogueTable (表 2) に記録する。なお、対話記録を保管するデータベースは Hermes 上で新規作成、アタッチ、デタッチが可能であり、利用するユーザの切り替えをスムーズに行える。

4.2. 要求

意図「要求」へ分類された場合、ユーザが Hermes に対して求めている行為が具体的に何であるか、ユーザの発言の主述語の語意によって分類する。

表 2 Dialogue Table

DateTime(datetime)	SG(text)	Utterance(text)	Classification(char(10))
2011/12/22 16:35	[sg_v200]f:1,自己...	自己PRでは学生時...	表明
2011/12/12 16:37	[sg_v200]f:1,会議...	会議には手帳を持っ...	表明
⋮	⋮	⋮	⋮

- PC 操作

主述語が「機械が運転を始める」概念 ID : 3d1110、「開いていたものを閉じること」概念 ID : 3cf5fa のような語意を持つ場合、ユーザの要求が行為「ソフトウェアを起動」「指定されたファイルを開く」「ソフトウェアを終了」である場合を PC 操作と判断する。

- スケジュール管理

主述語が「すでにある物に他の物を加えること」概念 ID : 3cf20f、「調査する」概念 ID : 30f874、「削除する」概念 ID : 0f314d、「変更する」概念 ID : 10819c のような語意を持つ場合、ユーザの要求が「予定の登録」「予定の検索」「予定の削除」「予定の変更」にあると判断する。後述する条件に加えて、ユーザの発言に「予定」「カレンダー」「スケジュール」「日程」のいずれかの語が含まれている必要がある。また予定の登録先として Google カレンダーを利用している。

- 天気予報

発話に「天気」というキーワードが含まれる場合、ユーザの要求が行為「天気予報情報の取得」であると判断する。行為「天気予報情報の取得」では、天気予報情報の取得に必要な要素を発言から取得し、live door 天気予報の xml ファイルを利用してユーザに天気予報情報を提示する。図 7 に、天気予報情報を取得する Hermes の実行画面を示す。

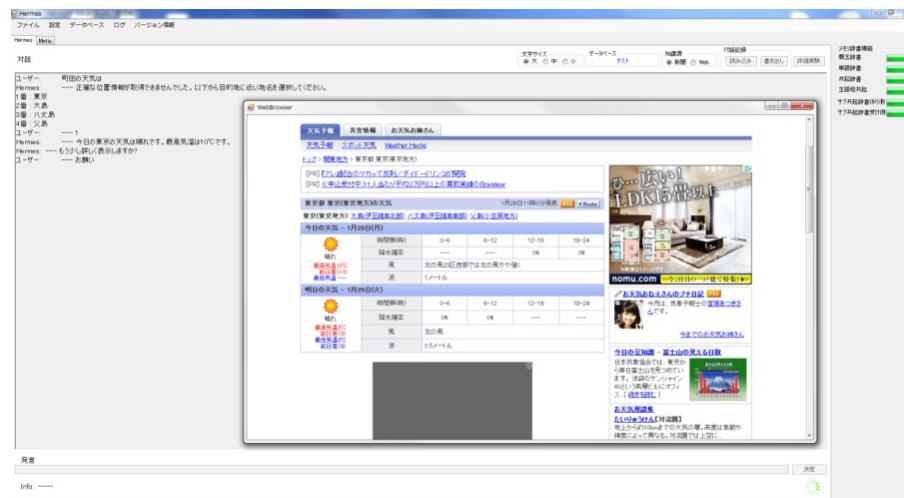


図 7 天気予報情報の取得

- 道案内

発話に「地図」「行き方」「案内」「ルート」というキーワードが含まれる場合、ユーザの要求が行為「道案内情報の取得」であると判断する。行為「道案内情報の取得」では、道案内情報の取得に必要な要素を発言から取得し、google map を利用してユーザに道案内情報を提示する。図 8 に、道案内情報を取得する Hermes の実行画面を示す。

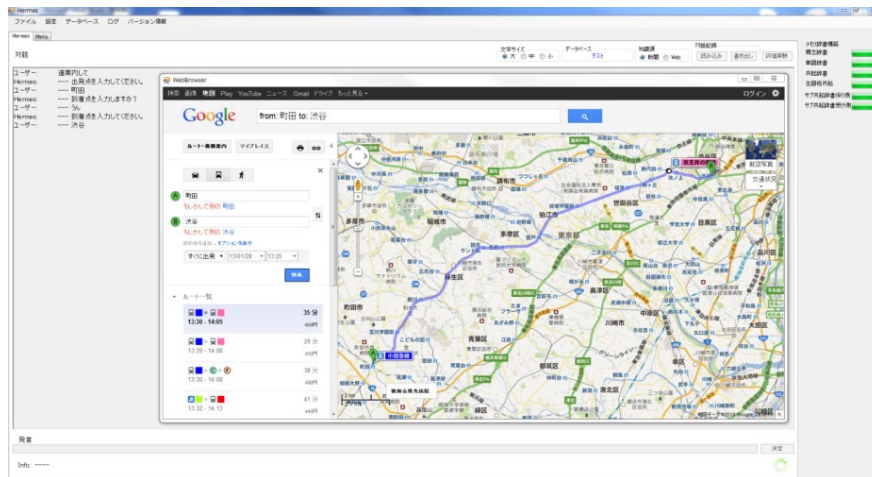


図 8 道案内情報の取得

4.3. その他の意図

- 挨拶

Hermes はユーザの発言の語意に応じた挨拶の言葉を返す。

- 肯定、否定

問い返しに対するユーザの回答を得る。意図「肯定」の場合、ユーザへ問い返した行為を実行する。意図「否定」の場合、ユーザへ問い返した行為を実行しない。

- 禁止、表明、叙述

意図「禁止」「表明」「叙述」においては現時点では簡単な応答を返す以外の行為は実装していない。Hermes は了解を表す簡単な応答を返す。

- その他

意図「その他」とは「相植など意図を持たないもの」であるので、ユーザの意図が「その他」であった場合には、ユーザは Hermes に対して具体的な行為を求めていると考えられる。そのため Hermes が行うのは簡単な応答を返すのみである。

4.4. 問い返し

Hermes は行為を実行するにあたって、その実行に必要な情報が不足している場合にユーザに対して問い返しを行い、不足している情報を対話的に取得する。問い返しには要素不十分型、選択肢型、イエスノー型の 3 種類を定義し実装した。以下、それぞれの問い返しについて説明する。

- イエスノー型

Hermes が行為を実行する際にユーザへ確認をとる必要がある場合に行う。行為の種類によってはその実行の際にユーザの理解を得ることが望ましい場合がある。例として、行為「予定の登録」における、予定の場所、予定の説明という任意の要素の取得が挙げられる。これらの要素は予定の登録において必ずしも必要ではなく、ユーザにこれらの要素を常に入力させることは望ましくない。そこでユーザに対してイエスノー型の問い返しを行い、行為の実行の可否を確認する。なお、ユーザの応答が可否のどちらであるかについては肯定、否定どちらに分類されるかによる。

- 選択肢型

Hermes が行為を実行する際にその対象が複数存在し、対象を特定できない場合に行う。行為の種類によってはその実行の対象が複数存在する場合があり、その特定が不可欠となる。例として、行為「指定されたファイルを開く」においてキーワードで対象のファイルを検索した結果、複数のファイルがマッチした場合が挙げられる。マッチしたすべてのファイルを開いてしまうとユーザが意図していない結果になることが考えられるので、ユーザに対して検索結果のうちどのファイルを開くのかについて選択肢型の問い返しを行い、開くファイルを特定する。なお、問い返しを行う際には検索結果に選択肢番号を振ってユーザに提示し、ユーザに開きたいファイルの番号を指定させる。ユーザの選択肢の指定方法は、「最初から三番まで」「1,2,4」「一と二と八」「すべて」などのような様々な表現に対応させる。

● 要素不十分型

Hermes が行為を実行する際に必要となる要素が不足している場合に行う。行為の種類によってはその実行において不可欠となる要素を持つ場合があり、その取得が不可欠となる。例として、行為「予定の登録」におけるタイトル、開始時間が挙げられる。この二つの要素は予定を登録する際に必須となっている。これらの要素をユーザの最初の発話から取得できなかった場合、各要素についてユーザに対して要素不十分型の問い返しを行い、タイトルや開始時間をユーザの応答から取得する。

5. 評価実験

モダリティと語意による話し手の意図理解の精度について評価実験を行った。実験対象として、青空文庫で公開されている小説から登場人物同士の対話を 200 件抽出した。発言ごとに本研究で定義した 9 種類の意図から最も適切と考えられるものを人手で選択し、それを正解データとした。実験は正解データと Hermes による意図理解の結果を比較して行い、正解データと Hermes の出力が一致したものを正解として集計した。

実験を行った結果、200 件の発言のうち 167 件の意図理解に正解し、全体としては 83.5%の精度を達成した。意図ごとの分類精度を表 3 に示す。

表 3 意図分類精度

質問	挨拶	肯定	否定	要求
81.4%	100.0%	80.0%	66.7%	90.2%
禁止	表明	叙述	その他	全体
50.0%	89.5%	22.2%	100.0%	83.5%

6. 他システムとの機能比較

表 4 に、しゃべってコンシェルと Siri との機能比較を示した。Hermes は、質問応答における回答精度が高い、対話記録を知識源とすることで過去の記憶を思い出すリマインダー機能がある、簡単な PC 操作を行えるという点で他のシステムに勝っている。その反面、電話・アラームといったスマートフォンアプリには未対応であり、実行可能行為拡充の余地がある。

表 4 他システムとの機能比較

	Hermes	コンシェル	Siri
スケジュール管理	○	○	○
質問応答	○	△	×
リマインダー	○	△	△
PC 操作	○	×	×
天気予報	○	○	○
道案内	○	○	×
メール、電話	×	○	○

7. 結論

本研究では、大きく二つの研究テーマを設定した。一つは、発話文から話し手の意図を理解し、それに応じた行為を実行する対話応答システム Hermes の開発であり、もう一つは、ユーザを支援する行為を実行することでユーザの役に立つシステムとして構築することであった。さらに、その中で対話記録を知識源として用いる事によって人手での知識追加を必要とせずにユーザの質問に答えることも目指した。

結果、ユーザの意図を 83.5%の精度で理解し、その意図に応じた行為を実行しユーザを支援するシステムとして構築できた。また、対話記録や新聞データ、WEB を知識源とした質問応答や、PC 操作やスケジュール管理、天気予報、道案内などのユーザの役に立つ行為を中心に実装した。

謝辞

本研究の一部は、(財)テレコム先端技術研究支援センターからの助成研究「言語で推論し音声で対話応答する携帯型秘書システムの開発」の研究助成金を基に行われました。当財団に深く感謝いたします。

参考文献

- [1] Weizenbaum, Joseph: "ELIZA - A Computer Program For the Study of Natural Language Communication Between Man And Machine", Communications of the ACM 9 (1) 36-45 (January, 1966)
- [2] Terry Winograd: "Procedures as a Representation for Data in a Computer Program for Understanding Natural Language", MIT AI Technical Report 235, (February 1971)
- [3] 後藤伸男, 中川侑美, 原田実: "自然語ルールと知識をベースとした対話システム Eve の開発", 情報処理学会研究報告, Vol.2009-NL-191, No.13, pp. 1-8, (2009.5)
- [4] 梅澤俊之, 加藤大知, 松田源立, 原田実: "意味解析システム SAGE の精度向上-モダリティと副詞節について-", 情報処理学会研究報告, Vol.2009-NL-191, No.4, pp. 1-8, (2009.5)
- [5] 原田実, 西岡晋太郎: "深層格とオントロジーを用いた回答抽出による質問応答システム Metis の回答精度向上", 情報処理学会第 73 回全国大会論文集, 1D-1, (2011.3)
- [6] 山田和正, 松田源立, 原田実: "照応解析システム ANA S Y S の精度向上-外界照応分離, 素性の洗練, ナイーブベイズ法の導入-", 情報処理学会研究報告, Vol.2010-NL-196, No.15, pp. 1-8, (2010.5)

- [7] 河原達也: "ここまできた音声認識技術", 情報処理, Vol.41, No.4, pp.436-439 (Apr.2000)
- [8] 河原達也, 松本裕治: "音声言語処理における頑健性", 情報処理, Vol.36, No.11, pp.1027-1032 (Nov.1995)