

曖昧な質問でも解決策を提案できる音声対話技術

Spoken Dialogue Technology to Understand Problems and Offer Solutions

永江 尚義

山崎 智弘

市村 由美

■ NAGAE Hisayoshi

■ YAMASAKI Tomohiro

■ ICHIMURA Yumi

パーソナルアシスタントと呼ばれるスマートフォン向けの音声対話アプリケーションの登場を契機に、決められたコマンドに制限されず自由なことばで話しかけられる音声対話技術に対する関心が高まっている。しかし、現在のパーソナルアシスタントには、してほしいことを明示的に話さないと十分な回答が得られないという課題がある。

東芝は、人と機械の更により自由で自然な対話を目指して、曖昧な内容でも意図を読み取って対応する対話技術を開発した。ユーザーがシステムに話しかけたとき、調べる内容や指示が明確でなくても文章中の単語や表現パターンから意図を推定し、あらかじめ大規模データから収集しておいた知識に基づいて解決策を提案できる。これにより、ユーザーが困り事や目的が曖昧な要望を話しかけると、人と相談したときのようなやり取りによってユーザーを解決に導くことができる。

Spoken dialogue systems such as personal assistant applications for smartphones have appeared in recent years. In order to hold a meaningful conversation with a conventional personal assistant application, however, it is necessary to input sentences containing explicit commands. Attention has therefore been increasingly focused on a spoken dialogue system to which users can speak freely, without the need for predetermined commands.

Toshiba has developed a spoken dialogue technology that can assist in resolving users' problems through more spontaneous human-machine dialogues. This technology makes it possible to provide adequate solutions to users through the estimation of intended meaning based on background knowledge collected from large amounts of data, including words and patterns in sentences, even when a user utters an ambiguous expression rather than giving a clear instruction to the system.

1 まえがき

パーソナルアシスタントと呼ばれるスマートフォン向けの音声対話アプリケーションの登場を契機に、音声対話技術に対する関心が高まっている。従来のコマンド型対話システムは、あらかじめ決められたコマンドしか受け付けることができなかったが、パーソナルアシスタントはある程度自由なユーザーのことばを受け付けられるため、より使いやすいユーザーインタフェースとなっている。

しかし、パーソナルアシスタントでも、「天気を教えて」のようにユーザーが“してほしいこと”を明示的に入力しなければならない。そのため、例えば空港で急病人を見つけて、ユーザーが「たいへんだ、急病人がいます」とシステムに相談しても、指示が不明確なため十分な回答は得られない。また、現在、コールセンターなどの顧客サポートの場面では人間のオペレーターが対応しているが、これらに対話システムが担うことができれば、24時間いつでも待ち時間なしですぐにつながる、満足度の高い顧客サポートを低コストで提供できるようになり、対話システムの応用範囲が広がっていくと期待される。そのためには、ユーザーの困り事や相談事のような潜在的な要望に応えることができる対話システムを実現する必要がある。このような困り事を相談できる対話システムを実現するには、「急病人

がいます」と話されたとき、それを“急病人のために何をすればいいか知りたい”という課題と捉え、その課題に対して、空港の緊急時の連絡先やAED（自動体外式除細動器）の設置場所を回答するといった解決策を提示することが必要である。

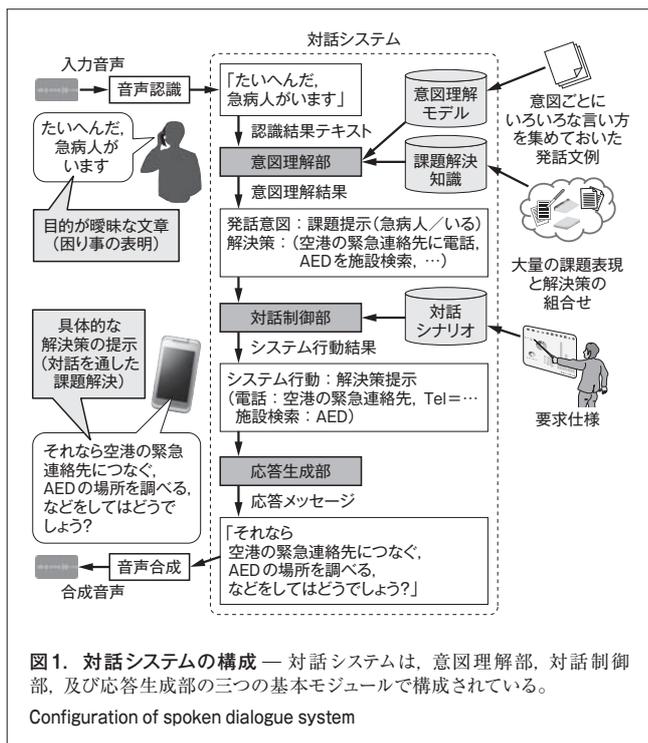
そこで東芝は、このような課題を解決できる対話システムの実現に向けて、ユーザーの様々な文章の意図を的確に把握し、内容から困り事などの課題を検出し、課題解決のための知識を活用して解決策を提案できる対話技術を開発した。ここでは、課題を解決する対話システムの概要と、課題解決のための知識の構築技術について述べる。

2 対話システムの概要

ここでは、課題を解決する対話システムの概要について述べる。

当社の対話システムの構成を図1に示す。対話システムは意図理解部、対話制御部、及び応答生成部の三つの基本モジュールで構成される。

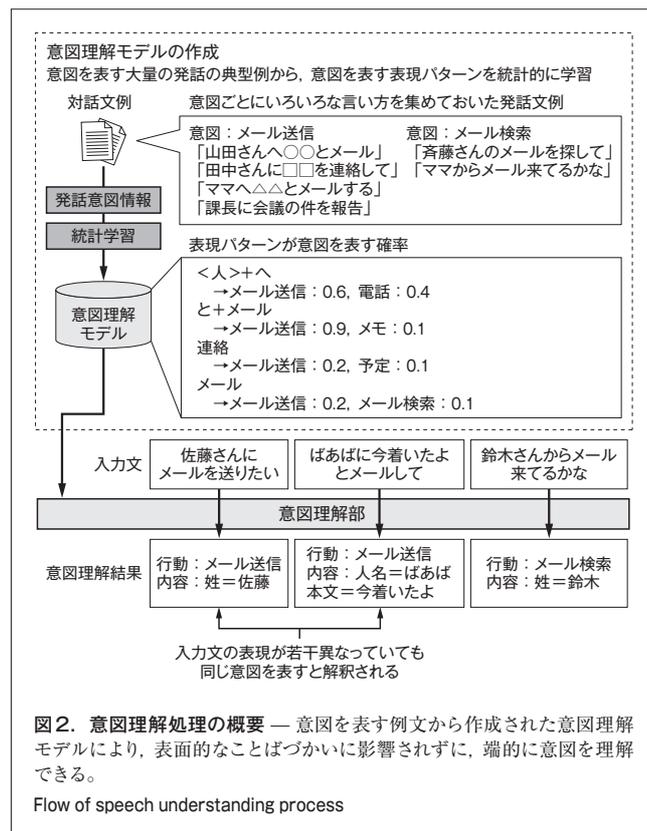
意図理解部はユーザーの入力文を解析して解釈し、発言の内容の意図を推定する。今回のシステムでは、機能実行の指示と課題の表明の二つを意図の解釈の対象とした。このような意図理解の処理は、単語や表現パターンがどのような意図



を表す傾向があるかをデータベース化した意図理解モデルに基づいて行われる。意図理解モデルは、入力される意図（例えば、メール送信）を伝えるときの様々な例文（例えば、「ママにメール」、「鈴木さんにメールを出して」、「今から出発しますとメールで連絡」）を収集し、意図を表すときによく使われる表現パターンとその確率を計算することで作成される。これにより、入力される文章で様々な表現が使われていても、表面的なことばづかいに影響されずに端的に意図を理解することが可能になる（図2）。

更に、意図理解部は、課題と解決策の組を表す課題解決知識を利用する。この知識は、入力された文の意図が課題の表明であるとき、その課題の解決策を提案するのに利用される対話の要となる知識である。この知識の詳細については3章で述べる。

対話制御部は、入力されたユーザーの意図と対話の流れを定めた対話シナリオに従って、対話システムの次の行動を決定する。対話シナリオには、様々な意図が入力されたときにそれぞれどのようなやり取りをするかという対話の流れが記載されている。例えば、メール送信（宛先：佐藤さん）という意図が入力されたとき、アドレス帳から宛先の人名でメールアドレスを検索し、sato@...というアドレスを見つけたら、send_mail(sato@...)というメールを送信するシステム行動を行う。ここで、宛先の名前がアドレス帳になかったときは、address_not_found（佐藤さん）というアドレス帳の検索失敗を通知するシステム行動を行う。対話シナリオは、構築するシステムに対する要求仕様に基づいて、システム設計者が定義する。



最後の応答生成部は、対話制御部が決定したシステム行動をユーザーに伝える応答メッセージを生成する。例えば、send_mail(sato@...)というメール送信のシステム行動のときは「佐藤さんにメールを送ります」、また、address_not_found（佐藤さん）なら、「佐藤さんという名前はアドレス帳にありません」というメッセージを生成する。

3 課題解決知識の構築技術

ここでは、曖昧な質問でも解決策を提案できる対話の実現に重要な役割を果たす、課題解決知識の構築技術について述べる。

3.1 目的情報抽出技術

人の様々な相談事に対して適切に応えるためには、大量の知識を用意する必要がある。そこで、知識を効率よく構築するために、当社が情報検索向けに研究を進めてきた目的情報抽出技術⁽¹⁾を応用して、対話システムが利用する課題解決知識を構築した。ここでは、この技術の概要を述べる。

目的情報抽出技術とは、Webテキストから個人の意見や経験を収集する手法の一つで、人の行動とその目的の関係を表す目的データを大量のテキストから自動的に抽出する技術である。目的データは、表1のように、“対象、動作、理由”の三つ組で構成される。例えば、“対象：清水寺、動作：行く、理由：紅葉がきれい”という三つ組は、“紅葉を楽しみたいので清水

表1. 目的データの例

Example of purpose data consisting of objects, actions, and causes

対象	動作	理由
清水寺	行く	紅葉がきれい
鍋	食べる	暖まる
キムチ鍋	食べる	辛い
ざっぱり	食べる	風邪(かぜ)
おかゆ	食べる	ざっぱり
レシピ検索	アプリケーション実行	娘の誕生日
スケジュール登録	アプリケーション実行	週末は旅行
病院	行く	風邪
コンビニ	行く	資料をコピーする

寺に行く”という人の行動と目的の関係を表している。このようなデータがあると、目的地を具体的に指定せずに漠然と「紅葉を楽しみたい」と指定されたときでも、理由に「紅葉」が出現する対象を調べることで、検索対象に清水寺や嵐山といった紅葉の名所を指定することが可能になる。このように検索時に対象や動作のキーワードが指定されていなくても、理由が指定されれば、関連する対象や動作を検索条件に加え、ユーザーが暗黙的に求めている要望に合致しているものを検索できるようになる。

3.2 スマートフォン向け課題解決知識の構築

今回、目的情報抽出技術を課題解決知識の構築に応用した⁽²⁾。この技術によって作られる目的データの“理由”を課題、“対象、動作”を解決策とみなすことで、対話システムが必要とする課題解決知識を表すことができる。この手法で収集された知識の有効性を検証するために、スマートフォン向けの対話システムを実際に構築した。

構築した課題解決の知識は、課題表現が格納される目的知識、スマートフォンで提供される機能のドメイン知識、及び“どのようなときにどのような機能を使うか”という目的知識とドメイン知識の関連付けを表す課題解決知識、の三つの知識で構成される(図3)。このデータから例えば、「コピーを忘れる」という課題に対しては、コンビニや文具店を地図で探すという解決策が得られる。

今回の対話システムの利用シーンとして、日常生活のなかで、Web検索や地図表示などのスマートフォンの機能を使ってユーザーを支援する場面を設定した。そこで様々な課題が入力されたときにシステムが適切な解決策を提案できるように、目的知識として、日常生活における人間の行動の理由や目的を表す課題表現をできるだけ広く収集することにした。表現を収集する対象として100億以上の文章から成る大規模コーパスを使用し、テキストから課題表現を自動的に抽出し、大量の目的知識を構築した。

ここで目的知識の構築には、単語の意味に注目する方法と、表現パターンに注目する方法を用いた。単語の意味に注目する

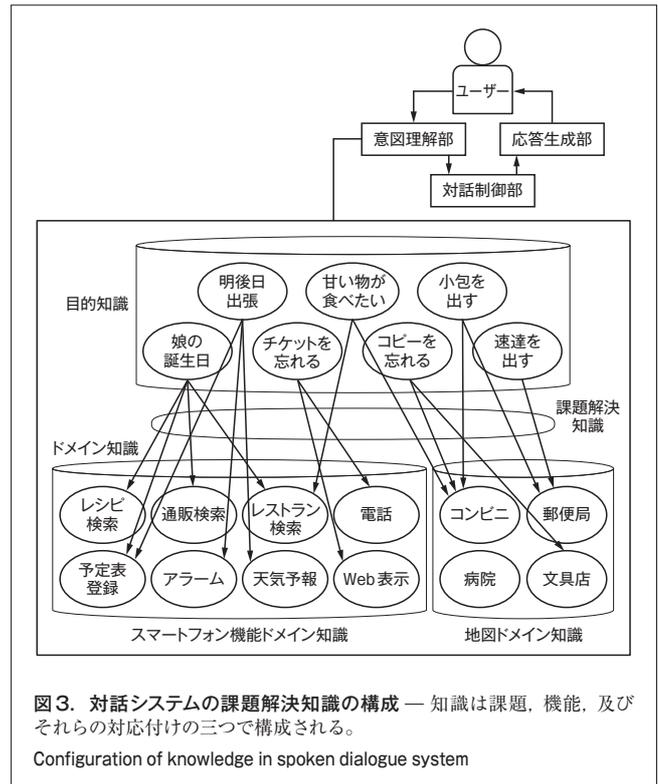


図3. 対話システムの課題解決知識の構成 — 知識は課題、機能、及びそれらの対応付けの三つで構成される。

Configuration of knowledge in spoken dialogue system

方法では、コーパスから行動のきっかけとなった事がらと人物や時間の関連を表す表現などを抽出した。“<人>+助詞+<イベント>”や“<時間>+助詞+<イベント>”といった表現に注目することで、「今度、夫が出張に行くんだけど」、「来年は入学式なので」などイベントがきっかけとなった課題表現が抽出された。一方、表現パターンに注目する方法では、人間の願望、人が置かれた状況、及び対象の状態を表す表現などをコーパスから抽出した。願望を表す表現として「甘いものが欲しいので」、状況を表す表現として「お給料が出たので」や「資料のコピーを忘れたから」、状態を形容する表現として「疲れたので」といった表現が抽出された。これら二つの方法で抽出された表現から低頻度な表現をフィルタリングして、日常生活における約4万件の課題表現を作成し、目的知識を構築した。

今回のスマートフォン向けのドメイン知識には、スマートフォンに搭載されている機能と、地図機能の中で検索される場所の2種類を用意した。スマートフォン機能ドメイン知識には、現在のスマートフォンに標準で搭載されている機能や、多くのユーザーが使用しているアプリケーションを登録した。地図ドメイン知識には、コンビニや郵便局など、ユーザーが地図に対して検索対象として指定すると想定される場所を表す語をリストアップして登録した。この結果、今回のシステムのドメイン知識は、約500件の機能と場所で構成されている。

これらの目的知識とドメイン知識の関連付けである課題解決知識は、課題が生じたときにスマートフォンのどのような機能で支援を行うべきかを表す知識となる。今回のシステムで

表2. 対話の例

Example of dialogue between user and spoken dialogue system

シーン設定：ビジネスマンが出先でスマートフォンを使って会議の場所を確認したときに、紙の資料の不足に気がついて、急ぎよコピーを取る必要ができてしまった

ユーザー	システム	説明
きょうのこれからの予定は？		スケジュール検索という指示と理解して、システムはスケジュール検索を実行
	“新製品検討会議 22A会議室（資料のコピーが6部必要）”という予定が見つかりました。	検索結果を表示
しまった、コピーを忘れた。		“コピーを忘れた”は課題と理解して、目的知識を検索
	コンビニ、インターネットカフェ、文具店などを地図で探してみるのはいかがでしょうか。	“コピーを忘れた”に関連付けられている“コピーを取れる場所を地図で探す”という解決策を提案
じゃあ、コンビニを見せて。		コンビニ検索を選択
	周辺のコンビニの地図を表示します。	現在地の地図を表示してコンビニをマークで表示 近くにあるコピーを取れる場所をすぐに見つけることができる

は、約4万件の目的知識と約500件のドメイン知識に対して、約25万件の関連付けを人手で構築した。関連付けには関連度も付与されており、課題に対する機能の使われやすさを表している。ここで三つの知識のうち、人間の行動の理由や目的を普遍的に表す目的知識は、対話システムのタスクやドメインとは独立な汎用的な知識である。そのため、新しい用途向けの対話システムを構築する際には、その用途向けのドメイン知識の設計を行ったうえで目的知識とドメイン知識の関連付けを構築すれば、様々な用途向けの課題解決知識を低コストで構築することが可能である。

3.3 知識を使った対話処理

システムは、ユーザーの意図が機能実行の指示の場合はその機能を実行する。一方、発話に具体的な指示が含まれていない曖昧な問いかけの場合には、課題の表明と判断し、目的知識を検索する。そして、該当する課題を検出した場合には、その課題に対応付けられているドメイン知識を解決策として提示する。この処理の流れを「コピーを忘れた」という入力が含まれる対話例で示したものが表2である。

まず、ユーザーが「きょうのこれからの予定は」と問いかけると、システムはスケジュール検索という意図を理解し、スケジュール帳に対してきょうの日付の予定の検索を実行したことを示している。スケジュール帳に資料のコピーの持参が必要な会議が登録されていたことに気づき、「コピーを忘れた」と話すと、この発話は具体的な機能実行の指示ではないため、システムは課題表明が入力されたと判断し、目的知識を検索したことを示している。目的知識の中に“コピーを忘れる”という課題表現があり、“コンビニを地図で検索”や“インターネットカフェを地図で検索”といった解決策と対応付けられていると、システムは“コピーを取ることができる場所を地図で検索する”という解決策を提案できる。このように、ユーザー自身

が具体的な解を持っていなくても、思ったことや困っていることをシステムに相談すれば、スマートフォンがみずからの知識に基づいてユーザーを支援する対話を行うことができる。

4 あとがき

人と機械の更に自由で自然な対話の実現を目指した、曖昧な内容でも意図を読み取って対応する対話技術について述べた。この技術により、ユーザーがシステムに話しかけたとき、調べる内容や指示が明確でなくても、文章中の単語や表現パターンから意図を推定し、大規模データから収集しておいた知識に基づいて、解決策を提案できるようになる。

また、この対話システムの実現に重要な役割を果たす課題を解決するための知識を、当社が情報検索向けに研究を進めてきた目的情報抽出技術を応用して構築した。人の意見や経験を“対象、動作、理由”の三つ組で表現し、大規模データから収集する目的情報抽出技術の手法を応用し、“理由”を対話で提示される課題、“対象、動作”を対話で提案すべき解決案とみなすことで、大規模データから知識を効率よく構築できた。この知識構築技術によって、様々な分野向けの課題を解決する知識を低コストで構築でき、対話システムを多様な応用に展開していくことが可能になる。

当社はこれからも、多彩な会話を行うための大規模な知識の構築や知識構築作業の自動化に取り組み、対話技術を活用した新しい製品やサービスを目指して研究開発を継続していく。

文献

- 若木裕美 他. “ロコミ情報からの目的情報抽出”. 第10回情報科学技術フォーラム (FIT2011). 函館, 2011-09, 情報処理学会; 電子情報通信学会. 第2分冊 p.15-22.
- 永江尚義 他. “課題解決知識を用いた音声対話検索システム”. 第11回情報科学技術フォーラム (FIT2012). 東京, 2012-09, 情報処理学会; 電子情報通信学会. 第2分冊 p.221-222.



永江 尚義 NAGAE Hisayoshi, D.Eng.

研究開発センター 知識メディアラボラトリー主任研究員, 博士 (工学). 音声言語処理技術の研究・開発に従事。情報処理学会会員。

Knowledge Media Lab.



山崎 智弘 YAMASAKI Tomohiro

研究開発センター 知識メディアラボラトリー研究主務。自然言語処理技術の開発に従事。電子情報通信学会会員。

Knowledge Media Lab.



市村 由美 ICHIMURA Yumi, D.Eng.

研究開発センター 知識メディアラボラトリー研究主務, 博士 (工学)。自然言語処理技術の研究・開発に従事。情報処理学会, 電子情報通信学会, 言語処理学会会員。

Knowledge Media Lab.