

顧客接点を高度化する サービスロボット基盤技術

Robot Platform Technology that Enhances Service Provider's Customer

● 上和田 徹 ● 今井 岳 ● 金岡利知 ● 武 理一郎

あらまし

近年、人と話をするコミュニケーションロボットを実用化し、ビジネスを立ち上げる動きが活発化している。たいていのコミュニケーションロボットは、ユーザーからの指示を受けてから動作するが、そのためにはまずユーザーが指示の仕方を覚える必要がある。そこで富士通研究所では、ロボットの方から能動的に話しかけることで役立つサービスへと導くロボットの実現を目指し、サービスロボット基盤技術を開発している。この基盤技術では、ロボット向けのアプリを多数用意しておく、その中からロボットが人の好みや状況に合ったアプリを選択してユーザーと会話する。このときのユーザーの反応は基盤側で収集し、以降のアプリ選択に反映していく。この基盤を適用したロボットによって、サービス提供者は様々なサービスを提示し、嗜好がマッチするユーザーとサービスとを結び付ける顧客接点の高度化が図れる。

本稿では、富士通研究所が開発した人とロボットとのインタラクションを通じてその人の嗜好や状態を獲得し、適切なサービスへと導くサービスロボット基盤技術について述べる。

Abstract

Communication robots that talk with people are being put to practical use as a means to start up businesses. Since most communication robots work only in response to the user's instructions, users have to learn the set of instructions for each robot. To overcome this problem, Fujitsu Laboratories is developing a platform technology that will enable robots to actively speak with users and introduce various useful services to them. With this platform technology, once many robot applications have been deployed, a robot can select the most appropriate one in accordance with the user's interests and concerns and the situation at the moment and use it to talk to the user. The user's responses are collected and used to optimize subsequent application selection. This robot platform enables service providers to enhance their customer contact points by better presenting various services that potentially match customer preferences. This paper explains Fujitsu Laboratories' robot platform technology for helping users choose appropriate services on the basis of their preferences and situation through interactions between the user and the robot.

ま え が き

人とICTの関係は、長らくコール&レスポンス方式で結ばれてきた。例えば、人が美味しいレストランを知るには、検索窓に「日本食 汐留」などとキーワードを打ち込んで指示すると（コール）、ICTはその指示に反応する文字列を返してくれる（レスポンス）。大量の情報や多種多様なサービスであふれかえる世の中では、存在すら知られていないサービスはコールされることもなく、人のICT利用は知識や経験のある範囲に閉じてしまうおそれがある。

しかし最近のスマートフォンでは、人が何も指示しなくても、「近所に渋滞があります」「現在の気温は10℃です」などと教えてくれる。このように人の状況（上述した例では場所）をセンシングし、その状況にふさわしい情報やサービスを推薦するということから、センス&リコメンド方式と呼ぶことができるであろう。ただし、このセンス&リコメンド方式にも課題がある。例えば、場所を知ることができても、頭が痛いことを知ることは難しい。この例からも分かるように、センシングできる内容には限界があり、推薦する情報が適切である保証はない。興味のない情報やサービスを提案され続けたら、ユーザーは迷惑と感じてしまう。

そこで筆者らは、対話を通じたアクティブなセンシングを行い、その結果に従って情報やサービスと人をつないでいくといった、いわゆるプロアクティブ・インタラクション技術の研究開発を行っている。例えば、頭痛かどうかをセンシングしたければ「頭痛がするの？」と人に聞いてしまえば良いと考えることができる。そしてこのような対話を、人と接するロボットという形態で広く提供することを目指し、そのようなロボットの基盤技術の研究開発に取り組んでいる。ロボットが人に話し掛けて対話しながらその人の嗜好や状況を理解していき、その人に合った情報やサービスへと導いていく。このようなロボットを通じてセンス&リコメンド方式の実現を目指している。

コミュニケーションロボットの技術動向

近年、人と話をするロボットを実用化する動き

が数多く現れ、ビジネスとしても立ち上がり始めている。その代表格がソフトバンク社のPepper⁽¹⁾である。Pepperは音声による対話のほか、首や腕、指の動き、胸に備えたタブレットなどの多様な入出力や表現の手段を備え、既に様々な商業施設に導入されている。ロボットの実現形態としては、Pepper本体側で主に動きやデバイスの制御を担い、音声認識など高度な情報処理はクラウドで行う。このようなクラウドを活用した構成は近年のロボットでも多く採用されており、今後の基本構造になると思われる。また、アプリ開発用のSDK（Software Development Kit）も公開されており、第三者がPepper向けアプリを開発・配布し、様々な用途に合わせて動作させることが可能である。現在、Pepper用アプリストアには100本以上のアプリが公開されている。

シャープ社のロボホン⁽²⁾も市販の人型ロボットである。こちらはAndroidスマートフォンとして作られており、持ち運びが可能なほど小型である。このロボホンもクラウドを活用した構成で、クラウド側のサービス機能が向上するのに伴って、ロボホンも進化する。更に、ロボホン向けのSDKも提供されており、ロボホン専用アプリを自由に開発し、Androidのアプリマーケットで配布できるようになっている。

このPepperやロボホンのような広く開かれたアプリエコシステムは、スマートフォンがそうであるように多数の開発者に支持され、新たな用途の開拓や利用シーンの拡大に大いに寄与している。

ロボットのユーザーインターフェースには、第一に音声期待されるが、音声インターフェース自体は既にスマートフォンでも一般的になってきている。Apple社のiPhoneには、音声での質問に答えたり検索したりすることのできるアプリSiri⁽³⁾が標準で搭載されている。Microsoft社も音声アシスタントCortana⁽⁴⁾をWindowsの標準機能としてパソコンに搭載し、普及を図っている。Google社も検索アプリで音声入力に対応しており、「OK, Google」に続く指示に基づいて検索を実行する⁽⁵⁾。更に、スマートフォンで収集できる操作や行動の履歴も使い、役立ちそうなアプリやサービスを推薦するGoogle Nowへと進化している。

このGoogle Nowは、センス&リコメンド方式を

指向しており、既に多くのスマートフォンユーザーに利用されている。ただし、何をセンシングできるかはスマートフォンでできる範囲に限定されてしまう。例えば帰宅後、スマートフォンを充電クレードルに置いたままではその間の行動は捉えられず、家の中では適切なリコメンドが得られないことになる。

そこで、家の中でも音声コマンドを気軽に使えるようにする新たな動きも急速に活発になってきている。Amazon Echo⁽⁶⁾はAmazon社のスピーカー型端末であり、備え付けのマイクで受けた音声をクラウド側で認識処理することで、Amazon社の物販や音楽配信などのサービスを音声で利用できるようにしたものである。この処理エンジンはAlexa⁽⁷⁾と名付けられ、音声コマンドで利用できる機能（スキル）を第三者が開発して後付けで搭載可能となっている。実際に、様々なサービス事業者が音声で自社サービス用のスキルを提供するという、新たなビジネスの枠組みが立ち上がっている。

Google社もGoogle Home⁽⁸⁾という家庭向けの音声インターフェース端末を提供している。こちらは家庭内の様々な機器を音声で操作するのが元々の狙いであったが、Alexa同様に幅広い応用の可能性を持つ。ただしAlexaもGoogle Homeも、受けたコマンドに応じて動くコール&レスポンス方式のサービス提供形態であり、ユーザーはコマンドを覚えて使いこなさなければならない。

サービスと人をつなぐコミュニケーションの実現

人の嗜好や状態を、対話を通じてアクティブにセンシングする方式は、GPSや加速度センサーなどによって人を外部からセンシングする従来の方法と比較して様々な利点がある。

まず、それはユーザーによる意識的な情報提供であり、知らないうちに位置情報を利用されていたといった問題を生じにくい。そして、センサーによる測定だけでは分かりにくい嗜好や体調についても、センシングできる可能性がある。更に、ユーザーによっては対話行為（とそれに呼応する情報やサービスの提案）そのものを楽しんでもらえる可能性がある。ロボットが日々新鮮な話題で、ときにはいくつかの質問も交えながら話し掛けることで、ユーザーの反応から好みやそのときの状況

などを理解していく。ユーザーを理解するにつれ、よりユーザーに合った話題が選ばれるようになり、ユーザーはロボットとの対話に常に関心を持てるようになる。ロボットとのコミュニケーションがネットサーフィンのようになり、ユーザーは様々なサービスを渡り歩きながら、関心のある情報や役立つサービスとの出会いを楽しむことになる。

しかし、実現に当たっては対話ならではの難しさも存在する。まず、嗜好という限りなく広いパラエティを持つものをセンシングするには、どのような対話を設計すれば良いのか。また、嗜好を探り当てたとして、その嗜好にマッチする情報やサービスがなかったらどうするのか。

筆者らはこれらの問題に対して、「情報・サービス自体に語らせる」という考え方で取り組んでいる。筆者らの対話システムのコンセプトを図-1に示す。ここでは、多種多様なロボット用のアプリ群が調停機能を通じて端末・ロボットにつながる様子が描かれている。

これらのアプリは、それぞれ対応する情報やサービスを持っている。例えば、ニュースアプリはニュースサイトにつながり、そのサイトに存在する情報をユーザーに伝えようとし、体重計アプリは何とかユーザーが体重計に乗るように促そうとする。

この方式により、筆者らは人の嗜好や状態に関する限りなく広い可能性の空間を当てなく探ることを避け、しかも飽きずに対話を続けてもらえる多様な問いかけを用意できると考えている。この方式に基づくロボットコミュニケーションは、ICTサービスの代理となる無数のアプリに支えられて実現されることになる。このような構成でロボットコミュニケーションを実現するための対話システムは、既存の対話システム⁽⁹⁾に比べて、次のような課題がある。

- (1) アプリ群を用いてセンス&リコメンドを行う一連の対話を生成する仕組み
- (2) アプリ数に対するスケラビリティと応答性
- (3) 人手によるアプリの開発コスト

これらを解決する技術については、次章で説明する。

上記に加え、「情報・サービス自体に語らせる」という考え方においては、ロボットという実体の

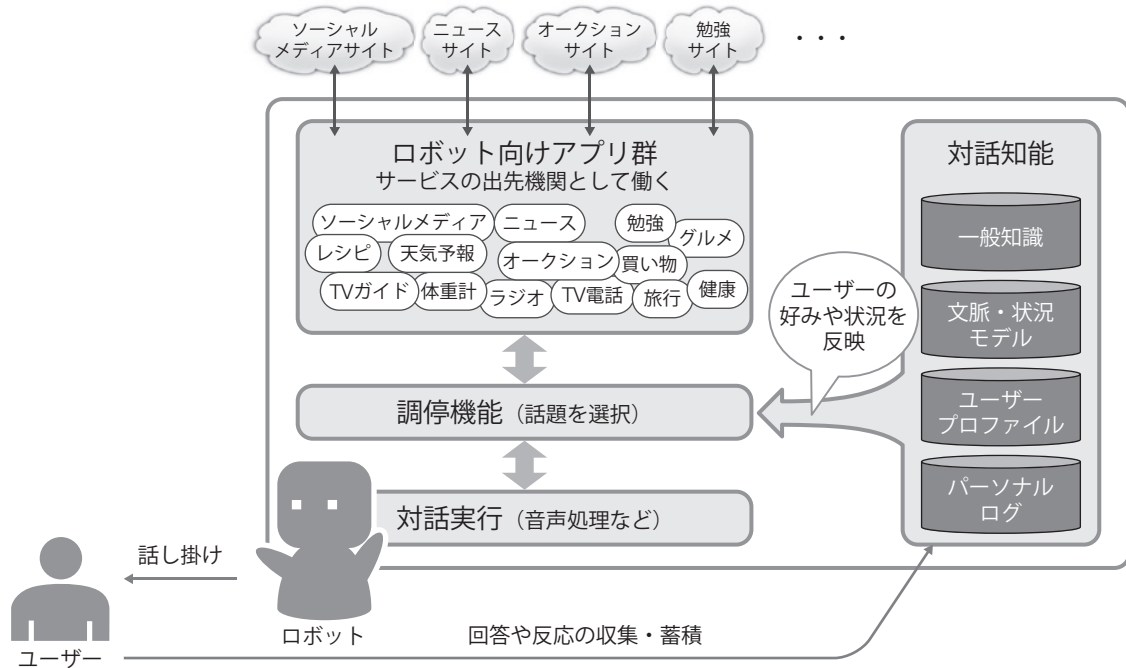


図-1 サービスと人をつなぐ対話技術のコンセプト

存在が非常に重要になってくる。ロボットのアプリを通じてサービスと人をつなぎ、情報を確実に伝えるためには、ロボットが人に対して働き掛ける意図を示し、相手に迷惑と感ぜさせないように、人の都合を把握して適切なタイミングを見計らわなければならない。

そこで、例えば何か伝えたいことがあるときには、ロボットがいきなり発話するのではなく、そわそわと小さく動き始めたり、人がこちらを向いたら話し始めたりするなど、人の行動を模した動作によって振る舞うようにする。こうすることによって、ユーザーの注意を自然に引き付けるとともに、その動きに対する人の反応をセンシングできる⁽¹⁰⁾。このような人を模した形状・動きを持つロボットが人と情報・サービスの間で介在することで、人を自然に対話に引き込むことができるようになり、人を煩わせることなくサービスを使ってもらえる機会を増やせると考えている。

また対人対話は、話す側と聞く側が交互に入れ替わる（話者交代する）連続的なやり取りで成立するものである。人は、相手に対する話し掛けの応答があることによって満足し、対話が促進される⁽¹¹⁾。応答は言語によるものに限られず、人は表情や視線の移動、うなずき、相づちなどの非言語

の振る舞いによって相手の話し掛けに対する理解を示す^{(12), (13)}。

近年、音声認識や対話処理など高度な情報処理がクラウドで行われるようになり、その精度が格段に向上した。一方で、通信遅延による結果取得のレスポンス低下が、話者交代における無言の間を発生させる原因になっている。この無言の間は、ロボットがこちらの発話を受け取れていないと感じさせてしまうことがあり、人に何度も発話させてしまったり、対話の興味を失わせたりすることになる。ロボットの非言語の振る舞いによる応答は、前述したように相手に理解を示し、対話を促進できるとともに、こうした無言の間を埋めることもできると考えている。

以上のような非言語の振る舞いを実現する技術についても、次章で説明する。

開発技術

本章では、ロボットが話す話題を取り扱う三つの技術と、非言語の振る舞いを実現する二つの技術について、それぞれ説明する。

● 話題の取り扱い

前章で述べた、多様なアプリを用いたセンス&リコメンド方式の対話を実現するための三つの課

題に対し、ロボットアプリ実行技術、対話サービス基盤技術、アプリ開発支援技術を開発した。これらの三つの技術が機能し、多種多様なアプリがこの基盤上で動作することにより、ユーザーの状況に応じた適切なサービスを提供できると考えている。以下に、それぞれの課題と開発技術について詳しく述べる。

(1) ロボットアプリ実行技術

対話によるセンス&リコmendを実現する基本的な考え方は、システムがユーザーに語り掛け、システムの発話内容に対するユーザーの反応に基づき、次のシステムの発話を生成することである。ユーザーの反応には、ユーザーの発話内容のほかに、ロボットがセンシング可能な人の動きや表情も含まれる。一般的に、シナリオに沿って書かれた対話のプログラムでは、多様なユーザーの反応に対するシステムの発話をあらかじめ用意することは難しい。既に述べたように、筆者らはこの課題に対して、ユーザーの多様な反応に対する発話を用意するのではなく、多様な情報やサービスを用意し、ユーザーの反応によって適したサービスを選択しながらそのときユーザーが求めるサービスをリコmendするというアプローチを採用することとした。

複数のアプリを渡り歩きながら対話を行うため、筆者らのシステムでは、アプリの構造を小さな実行ブロックに分割し、実行ブロック単位でアプリを切り替えながら対話を進めていくこととした。ここで、対話の一応答を実行ブロックとし、「対話ブロック」と呼ぶ。各アプリは複数の対話ブロッ

クで構成され、複数の応答による対話も可能である。ユーザーの反応が想定から外れた場合や、より適した発話が存在する場合には、別の対話ブロックが実行されることもある。

一方で、一つの応答で終了する、自身のアプリへの呼び水となるような雑談の対話ブロックを用意し、その反応をセンスすることにより、適切なタイミングで自身のサービスをリコmendできる(図-2)。このような仕組みを持つことにより、よく練られた対話シナリオや、特定シーンにおける大量の対話例からの学習に基づいて生成された対話に比べて不自然になるかもしれないが、インタラクティブにユーザーの反応を引き出し、それに応えていくことができる。

現在のシステムでは、ユーザーの反応に合わせて対話ブロックを選択するために、アプリ開発者が対話ブロックに対して付与する関連キーワードや、対話ブロックの発話文に含まれる語と直前の応答、ユーザー発話内容、ユーザーの嗜好などとの一致度合いを求め、その結果に応じて対話ブロック選択を判断している。

(2) 対話サービス基盤技術

ここまで主に、センス&リコmendの実現方法について述べた。実際にこの仕組みが人とICTサービスをつなぐ役割を果たすためには、膨大な数のアプリがこの仕組みの上で動作する必要がある。従来の対話システムでは、ユーザー数に対するスケーラビリティは検討されてきた。⁽¹⁴⁾ 本システムでは、アプリの数に対するスケーラビリティも考慮する必要があり、アプリの数やユーザーの数が増

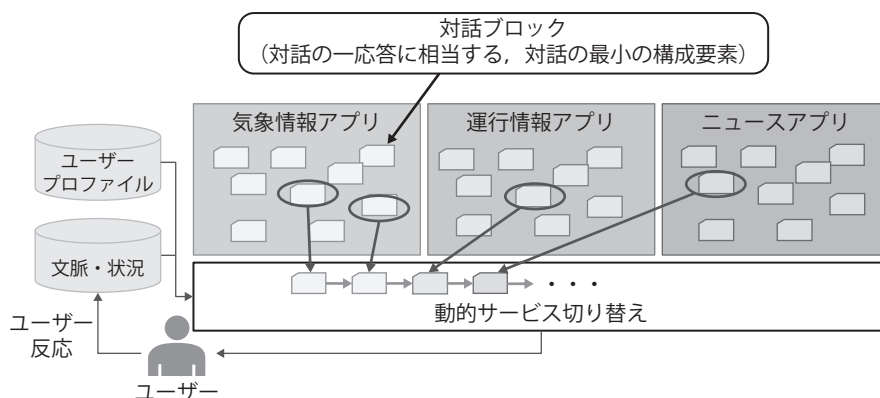


図-2 複数アプリを切り替える対話の仕組み

加しても対話として成立するように、一定の応答性を保ちながら動作を継続できることが求められる。このため、筆者らが開発した対話サービス基盤では図-3のような構成を採っている。

アプリは、対話サービス基盤上で動作するのではなく、対話サービス基盤のリクエストに応じて対話ブロック群を提供する役割を果たす。対話サービス基盤は、事前にアプリから対話ブロックを取得して保持している。また、ユーザーの発話内容を解析・解釈して、対話状態、文脈情報として保持する。これらに基づいて対話ブロックを選択・実行し、発話候補の対話ブロックから文脈に適した対話ブロックを選択して、ユーザーに対する発話を行う。事前に対話ブロックを読み込んでおくため、ユーザーの発話に対する一連の流れの中で多数のアプリに対する処理やアクセスは発生せず、一定の応答性を保つことができる。

一方で、対話状態の変化に応じてアプリ選択が行われ、発話候補の対話ブロック群を動的に変更する。文脈情報に応じて発話を選択する応答性の高いループと、発話候補の対話ブロックを変更する比較的応答性の低いループを同時に動作させることにより、応答性の確保と膨大な数のアプリからの発話選択を両立している。また、計算量の観点から見ると、ユーザーごとに各時刻で考慮すべき対話ブロック数を限定しない従来の方法では、システム全体としてアプリ数×アクティブなユーザー数のオーダーの計算量が必要となる。これに

対し、非同期に段階的な選択を行う本技術では、アクティブなユーザー数のオーダーの計算量に抑えることができる。

(3) 本基盤向けアプリ開発支援

これまで述べたとおり、本システムはアプリの数が膨大になってくると、基盤としての効果が顕著になる。ユーザーは、多様なアプリにより様々な側面でセンス&リコメンド方式のサービス提供を受けることができる。一方、ICTサービスの提供者は、本基盤向けのアプリを用意することにより、ユーザーの状況に応じて適切にサービスを提供できる。しかし、新たな基盤向けにアプリを新規に開発するコストが大きいと、サービス提供者が参入する障壁になってしまう。

そこで本基盤向けアプリの開発を支援するために、極力既存のWebコンテンツやWeb API (Application Programming Interface) を活用して、少ない工数でアプリを開発できるツールを提供する。Webコンテンツのような画面上のレイアウト構造を持つアプリを対話用のアプリにする際には、レイアウトの構造をどのように対話の構造に対応付けるかが大きな問題となる。そのため、特定のレイアウト構造を抽出して、対話の構造にマッピングすることを支援するツールを提供する。現状は、プロトタイプとしてWebコンテンツ上の構造 (DOM: Document Object Model) に対して、対話の構造のテンプレートを割り当てることにより、対話のコンテンツを出力できる機能を提供している。

テンプレートには、リストを選択する、メニューを読み上げる、RSSフィードを用いて新着情報を伝えるなどの機能を用意している。今後もこれらのテンプレートを拡張し、更に既存コンテンツの構造の種類に応じたアプリへの変換方法を体系化し、容易にアプリを生成する開発支援ツールを開発する予定である。対話アプリを一から開発する場合には、一般に全ての対話のシーケンスとそれに対する応答を考慮した設計を行い、実装したものに對するチューニングを繰り返すことになる。本開発支援ツールでは、設計と実装の手間を大幅に減らし、初期導入のコストを全て人手で対話シナリオを開発する場合に比べて、10%程度に抑えることを目指している。導入後は、Webのコンテ

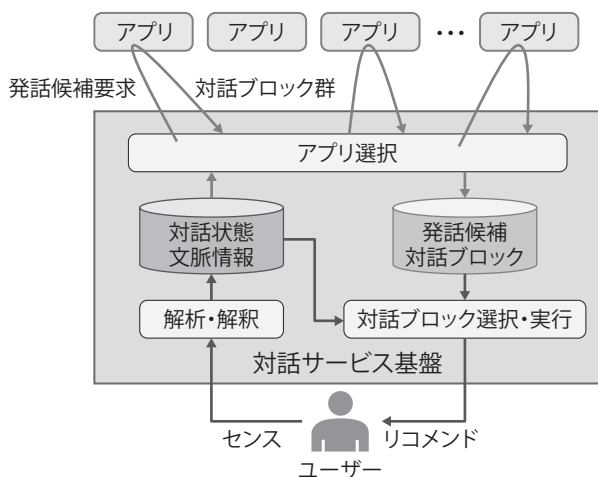


図-3 対話サービス基盤の構成

ンツから自動的にアプリを生成するため、Webコンテンツの構造に大きな変更がない限り、その内容に変更があっても変更するコストはかからない。

● 非言語の振る舞い

次に、非言語の振る舞いを実現する技術について説明する。アプリとユーザーをつなぐロボットにおける音声対話制御技術として、対話開始タイミングをセンサー情報に基づいて決定する対話開始タイミング制御技術と、話者交代時の無音時間を削減する話者交代制御技術を開発した。ここで、簡単に対話情報のフローについて触れておく。

ロボットシステムのブロック図を図-4に示す。クラウド上にある対話サービス基盤からプッシュされた情報は、外部接続インターフェースを経て音声対話制御部で処理される。こうして、ユーザーの状況の人検出・認証情報に基づいて対話開始タイミングを制御することにより、タイミングを計って音声合成しテキストを音声データに変換して発話する。一方、ユーザーからの返答は、音声検出・認識を経て音声対話制御に入力され、話者交代制御によりユーザーに発話を受け付けたことを返答し、外部接続インターフェースを経由して、対話サービス基盤にユーザーの発話情報が送られる。

(1) 対話開始タイミング制御技術

対話サービス基盤にユーザーの発話情報が送られたとき、ロボット側では、ユーザーの有無やロボットに対する注意などのインタラクションの状

況に基づいて、アプリからの働き掛けのタイミングを計る。この対話開始タイミング制御では、カメラを用いた人検出や音声検出などのセンサーから得られる情報により、ユーザーとロボットとのインタラクションの状況を把握する。このインタラクションの状況に基づいて、ユーザーがロボットに注意を払っていないときには、ロボットは目配せやそわそわするといった非言語の振る舞いを行うことでユーザーの注意を引き付け、アプリからの働き掛けを行う。

このロボットの制御技術には、「子ぐま型ソーシャルロボット」⁽¹⁵⁾で開発した親和的インタラクションに基づくサービスを実現する階層的アーキテクチャー⁽¹⁶⁾を用いている。このアーキテクチャーにより、生物感を表現する自律的な動作とアプリからの働き掛け動作を自然に重畳している。そして、ユーザーの注意を引き付ける非言語振る舞いの要素検証において、その振る舞いがないときと比較して、人がロボットからの視線を受けていると意識する割合を約60%向上させる結果が得られている⁽¹⁰⁾。

(2) 話者交代制御技術

対人対話は、話す側と聞く側が交互に入れ替わる連続的なやり取りで成立するものである。相手の発話に対して応答を示すことで相手に満足感を与え、対話が促進されていく。この対話を支えているのが非言語の振る舞いであり、人はそうした振る舞いを応答時に巧みに利用することで、実質的な対話を途切れさせずリズム感の良いやり取りを続けている。筆者らが開発した話者交代制御技術では、対話を途切れさせずリズム感の良いやり取りを実現するために、次の三つの施策を行っている。

一つ目は、ユーザーが発話する音声情報と発話内容に基づいて、ユーザーの発話における切れ目を検出し、「うんうん」のような相づちを挿入する。これにより、ロボットが発話を受け入れていることをユーザーに即座に通知する効果をもたらすと同時に、無言の間を作らないようにできる。更に、ロボットのように動きが表現できるデバイスであれば、うなずきや視線移動のような非言語動作を組み合わせることで、より一層自然に見せることができる。

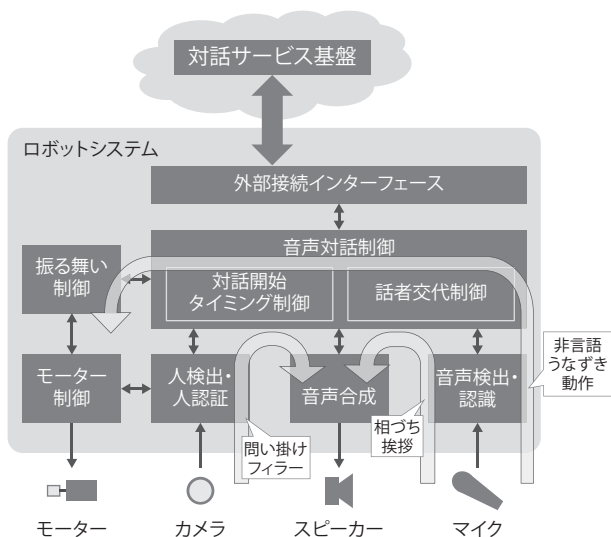


図-4 ロボットシステムのブロック図

二つ目は、ロボットからの発話が長くなったときに、ユーザーの注意を引き付けてから、確実に情報を伝えるために、「あのお～」や「え～っと」のようなフィラーを発話開始に挿入している。⁽¹⁰⁾

三つ目は、あらかじめユーザーが発話しそうな単語をロボット本体に搭載し、ローカルで音声認識を行っている。このようにローカル処理を行うことで、挨拶や短い応答に対して、クラウド上で行われる音声認識処理や言語処理に時間をかけず即座にユーザーの応答に対して反応できる。

これらの音声インターフェースの施策によって、ユーザーの発話終了から、システムがそのユーザーの発話を認識して、次のロボットの発話を行うまでに発生する2～3秒の無音時間に対応する。こうして、ユーザーの発話に対する即時応答とロボット発話前のフィラーなどにより、実質的にユーザーが体感する無音時間を1秒未満に短縮することで、リズム感の良いやり取りを実現している。

本技術の適用例

対話を主体とするサービスロボットは、ユーザーに音声で情報を伝える音声発話や、ユーザーからの発話を受け入れる音声認識といった機能要素だけでなく、対人対話を基にした対話を促進する必要がある。このため、ロボットの顔や形といった擬人的表現も重要な構成要素である。更に、アイコンタクトやうなずきといった非言語の振る舞いを実現する機構は、言語による対話を補完し、かつユーザーの注意をさりげなく引き付けるための重要な要素である。このため、多くのロボットが最低限の機構として頭部の動作機構を取り入れている。

こうした擬人的なロボットを利用することで、ユーザーにロボットが社会的な存在であると印象付けることを助け、その先にあるシステムからユーザーへの働き掛けをスムーズにする役割を担う。同時に、音声対話におけるロボットの振る舞いがユーザーとの対話促進に貢献する。以下に、ロボットを介して本システムにより生成されるセンス&リコメンド方式の対話の例を示す。ここで、S1からS8はシステムの発話を表し、U1からU8はユーザーの発話を表す。

● 対話例1

S1：川崎に新しいラーメン屋さんができみたいだね。(レストラン検索アプリ)

U1：え、ほんとに？詳しく教えて。

S2：XX屋。先週オープンした濃厚豚骨ラーメンのお店だよ。(レストラン検索アプリ)

U2：そうかあ、また今度行こうかな。

S3：地図を出しますか？(レストラン検索アプリ)

U3：いや、今はいらぬ。

S4：ラーメンといえば、そろそろお腹がすきませんか？ご飯を作りましょう。(レシピ検索アプリ)

U4：そうだね。今日は豚肉を使わないと。

S5：豚肉で検索しますか？(レシピ検索アプリ)

U5：お願い。

S6：チンジャオロースはどうですか？(レシピ検索アプリ)

U6：ピーマンがないなあ。

S7：ピーマンで検索しますか？(レシピ検索アプリ)

U7：いや、しないよ。

S8：ピーマンならYYスーパーで今日は特価だよ。(YYスーパーアプリ)

U8：買いに行こうかな。

この例では、これまでの対話の履歴やユーザーの嗜好情報に基づき、発話候補の対話ブロックがあらかじめ取得されている。S1では、ラーメンが好きというユーザーの嗜好情報に基づき、レストラン検索アプリが提供する新規オープンの店舗の情報が選択される。S2では、ユーザーの発話に反応して詳細情報を発話し、S3で更に追加の情報を発話しようとする。しかし、U3により拒否されたため、S4では食事で文脈の近いレシピ検索アプリが提供する対話ブロックが選択されている。S4からS6ではレシピ検索アプリが料理を提案している。U6のユーザー発話に対して、レシピアプリがピーマンに対して可能な応答としてS7を発話するが、これは的外れであり、U7でユーザーに否定されている。次に、YYスーパーのアプリの対話ブロックが選択され、S8を発話している。このように、ユーザーの反応を見ながらセンス&リコメンド対話が生成される。

ユーザーがこのシステムを使い始めた段階では、対話例1のようなトライ&エラーの会話を行うが、その後、ロボットはユーザーとの対話を繰り返す

ことにより、ユーザーの個人的な情報や、ユーザーの会話の傾向や嗜好などの情報を得ていく。こうして得た情報を活用して、個人的な思い出や出来事に関連する話題を投げかけ、ユーザーの反応をセンシングすることにより、ユーザーにとって思いがけない提案ができるようになると考えている。対話例2に、その一例を示す。

● 対話例2

S1：そういえば、去年の今頃は家族でJリーグの観戦に行ったよね。(基盤の話題提供機能)

U1：そういえば、川崎フロンターレの応援に行ったね。

S2：川崎フロンターレのスケジュールを調べますか？(Jリーグアプリ)

U2：いや、いいよ。子供たちそんなに喜んでなかったからなあ。やっぱり公園とかの方がいいみたいだな。

S3：XX公園は先週行ったよね。(基盤の話題提供機能)

U3：そうだなあ。たまには違うところに連れて行こうかな。YY公園とかがいいかな。

S4：YYといえば、おいしいパスタ屋さんがあります。(外食検索アプリ)

U4：では、お昼はそこにしよう。

この例では、S1で基盤の機能により、過去のユーザーの行き先の情報に基づいた話題を提供する。一般的な話題より、特定の事象に関する話題の方がユーザーの言葉を引き出しやすい場合が多いからである。この例では、それに反応してU1で特定のチーム名をユーザーが発話し、そのチーム名に反応したJリーグアプリがS2で対話を進めようとする。しかしU2でユーザーはネガティブな反応をし、新たに公園という単語を発話する。家庭内の会話で出現する「公園」という単語は、一般名詞よりも近所の特定の公園を指すことが多いと考えられる。そのためS3では、基盤の機能により、普段よく行くXX公園について、過去の対話から把握している事実を述べている。U3では、ユーザーからYY公園という言葉が引き出されたため、S4では外食検索アプリによって、YYという地区にあるお店を紹介できた。

この例のように、一般的な話題だけでなく個人的な事象に関する話題を提供することにより、ユーザー

の意思を反映した豊富な情報をユーザーの発話から引き出せる可能性があり、より適切なサービスを提案できる。

む す び

本稿では、ロボットを通じてセンス&リコメンド方式で様々な情報やサービスを人に届ける基盤となる、サービスロボットにおけるプロアクティブ・インタラクション技術について述べた。

サービス提供者がこの基盤上で動作するアプリを用意することで、ロボットと人との対話の中にサービスに導く会話が挿入され、新たな顧客との接点を生む機会となる。一方ユーザーにとっては、新鮮かつ自分の関心に合う話題が提供され、それが自分に有用なサービスとの出会いとなる。この利点がユーザーに受容されると、様々な嗜好データを得られるようになり、データ活用ビジネスに展開できるようになる。ただし、ロボットの見た目は人によって好みが異なったり、その時点の流行に評価が左右されたりするものである。そのため、場面や用途に応じて自由にロボットの姿を選べるよう、この基盤技術はロボットの見た目について特に規定していない。この基盤を活用することで、サービス提供者はデータの扱いは本基盤に任せ、ロボットの外見をユーザー層にフィットさせることに注力できる。

今後、この技術を様々な用途に適用し、幅広い利用シーンで顧客接点の高度化を進めるとともに、それによるデータビジネスの拡大を目指す。

参考文献

- (1) ソフトバンク：Pepper.
<http://www.softbank.jp/robot/consumer/>
- (2) シャープ：ロボホン.
<https://robohon.com/>
- (3) Apple：Siri.
<http://www.apple.com/jp/ios/siri/>
- (4) Microsoft：Cortana.
<https://support.microsoft.com/ja-jp/help/17214/windows-10-what-is>
- (5) Google：ウェブ検索.
<https://www.google.co.jp/search/about/>
- (6) Amazon：Amazon Echo & Echo Dot.

<https://developer.amazon.com/echo>

(7) Amazon : Alexa.

<https://developer.amazon.com/alexa>

(8) Google : Google Home.

<https://madeby.google.com/home/>

(9) 河原達也：音声対話システムの進化と淘汰－歴史と最近の技術動向－. 人工知能学会誌 28 (1), 2013.

(10) 金岡利知ほか：親和的インタラクションに基づく人に優しい情報提供. 第26回人工知能学会全国大会予稿集, 2012.

(11) 大坊郁夫：しぐさのコミュニケーション－人は親しみをどう伝えあうか－. サイエンス社, 1998.

(12) S. Andrist et al. : Conversational Gaze Aversion for Humanlike Robots, In Proc. HRI '14, 2014.

(13) A. Terrel et al. : A Regression-based Approach to Modeling Addressee Backchannels, In Proc. SIGDIAL '12, 2012.

(14) 荒金洋介ほか：音声対話システムにおけるスケーラビリティ評価モデルの検討. 情報処理学会論文誌 46 (9), 2005.

(15) 人にやさしい端末「子ぐま型ソーシャルロボット」ユーザーとの親和的な関係を築くインタラクション技術. 富士通ジャーナル2010年4月号, 2010.

(16) 今井 岳ほか：親和的インタラクションに基づくサービスを提供する子ぐま型ソーシャルロボットの開発. 第12回計測自動制御学会システム・インテグレーション部門講演会予稿集, 1C2-2, 2011.



金岡利知 (かなおか としかず)

グローバルビジネス戦略本部
サービスプラットフォーム戦略企画室
ロボティクスサービス企画開発部
ヒューマンロボットインタラクションに関する研究・開発に従事。



武 理一郎 (たけ りいちろう)

IoTシステム研究所
ヒューマンセントリックコンピューティング全般の研究開発と事業化に従事。

著者紹介



上和田 徹 (かみわだ とおる)

グローバルビジネス戦略本部
サービスプラットフォーム戦略企画室
ロボティクスサービス企画開発部
サービスロボット向けのインタラクション技術の研究開発に従事。



今井 岳 (いまい たかし)

フロントテクノロジー研究所
UXソリューションプロジェクト
分散協調システムおよびロボットサービスプラットフォームの研究に従事。