

会話連動ナビゲーションシステム Information navigation system based on conversation

向垣内 岳弥† 高田 慎也† 横関 大子郎† 酒井理江† 村山 隆彦† 新井 克也†
MUKAIGAITO Takeya TAKADA Shinya YOKOZEKI Daigoro SAKAI Rie MURAYAMA Takahiko ARAI Katsuya

1. はじめに

コミュニケーションを行う際に、表面的に伝達される情報だけではなく、それらに関連のある情報を参照したり、相手の意図や感情、ユーザが置かれている状況といった背景情報(コンテキスト)を考慮して対応する場合がある。例えば、話題に関連のある情報の検索を行ったり、相手との過去のメールのやり取りや自分の予定表を見ながら電話をすることや、相手が急いでいる場合は要点だけを述べるといったことを行っている。これらの行為は、会話を行う際に不足している情報を、ユーザが自ら補う行為と言える。

しかし、会話を行っているときに、ユーザが自ら会話に関連のある情報を探すことは、ユーザにとって負担であり、会話の進行を阻害する要因になる。関連情報を探すためには、現状、キーボードやマウスを操作する必要があり、その間、会話が遮られることになる。このため、会話に関連のある情報を自動的に提供することができれば、ユーザに負担を与えることなく、コミュニケーションをより円滑に進めることができ、コミュニケーションの内容を充実したものにできると考えられる。

近年、音声認識技術[1]や音声からの感情認識技術など、音声に対する認識技術の進展が著しい。本稿では、これらの認識技術と、会話からの特徴的なキーワード抽出技術、コンテキスト抽出技術を組み合わせることで、会話に関連する情報をリアルタイムに提供する、会話連動ナビゲーションシステムのコンセプトを紹介し、その実現方式について述べる。

2. 関連研究

システムとユーザとの音声対話によって、適切な関連情報をナビゲートする対話システムの研究は、これまでに多数行われている。例えば、電話を通じて利用可能なサービスが商品化されている[2]。また、大前らによる音声情報案内システムにおける発話感情理解に関する検討[3]等がある。

ユーザ同士の会話に対する情報支援システムとしては、例えば、後藤は、話者が通常の会話に現れないような特殊な発話を行うことで、支援を受けたい部分をシステムに識別させる音声入力方法を提案している[4]。また、城塚らは、専用TV会議システム上で、コンサルタントが行う相談業務を支援するシステムを提案している[5]。

本稿では、ユーザ同士が電話上で行う通常の会話に着目し、会話音声や会話音声から得られる感情、ユーザの特性を基に、会話に応じた情報をユーザに意識させることなくリアルタイムに提供する、会話連動ナビゲーションシステムを提案する。

3. 会話に対する提供情報

電話で会話中のユーザに対して本システムが提供する情報を示す。提供する情報は、会話内容に関する(1)詳細情報と、(2)連想情報とに分類される。それぞれの情報は、さらに、a)ユーザに依存しない一般的な情報と、b)ユーザに依存した情報とに分類される。

(1) 詳細情報：会話の内容に関するより深い情報

会話の内容に関して、ユーザがより詳しく知りたいであろう情報を提供する。ユーザが必要とする情報を比較的明確にイメージしている場合に有用である。

a)ユーザに依存しない一般的な詳細情報の例として、会社や商品等の固有名詞に関連する、公式ウェブページや、ユーザが知らない単語の意味などが挙げられる。

b)ユーザに依存した詳細情報としては、例えば、「新宿」と「8時」という言葉に対して、現在ユーザがいる場所から、新宿に8時に到着するための電車の時刻表の提供がある。

(2) 連想情報：会話の内容に関するより広い情報

会話の内容に関して、関連のあるキーワードや新たな話題を提供する。ユーザが必要とする情報を、すぐには思い付かない場合に有用である。

a)ユーザに依存しない一般的な連想情報の例として、「スポーツ」のような一般的な名詞から、「野球」や「サッカー」といった、より具体的な内容の提供が挙げられる。

b)ユーザに依存した連想情報の例としては、上記の場合において、ウェブ等へのアクセス履歴を基に、「野球」ではなくそのユーザがより興味を持っている「バレーボール」を提供することが考えられる。

上記のような情報を、会話を遮ることなく、また、ユーザに負担を与えることもなく、電話での会話中にリアルタイムに提供することにより、コミュニケーションをより円滑に進めることができ、コミュニケーションの内容を充実したものにできると考えられる。例えば、待ち合わせ場所や時間を決める際に、会話中に地図や時刻表が表示されるウェブページを手作業で閲覧・検索していたものが、会話に応じて自動的にユーザの端末に表示されることで、本来の会話には関連しない作業を省くことができ、コミュニケーションがより円滑に進むと期待される。また、「スポーツ」から「野球」や「サッカー」といった連想情報を提示することで、会話内容を膨らます効果が期待され、コミュニケーションがより充実したものになると考えている。このように、会話に連動して、詳細情報や連想情報を提供することは有効であると考えられる。

† 日本電信電話株式会社
NTT 情報流通プラットフォーム研究所

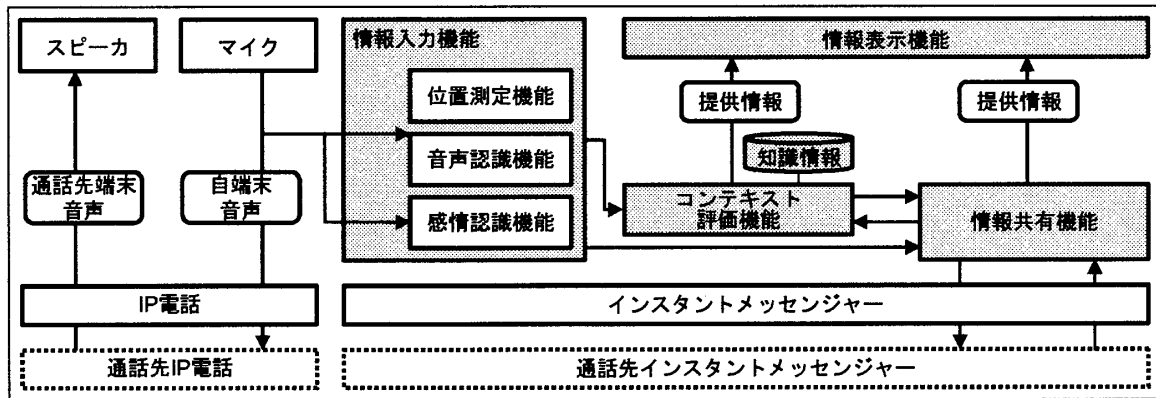


図1：会話連動ナビゲーションシステムの構成

4. システム構成

提案する会話連動ナビゲーションシステムの構成を図1に示す。提案システムは、情報入力機能、コンテキスト評価機能、知識情報、情報表示機能、情報共有機能から構成される。

情報入力機能は、音声認識機能や感情認識機能、位置測定機能等、ユーザの情報を取得する機能から構成される。ここで、電話を通じて交換される会話の内容やユーザの感情、ユーザが現在いる位置情報等のユーザに関する基本的な情報を取得する。

コンテキスト評価機能は、情報入力機能で取得した各種情報と情報共有機能で取得した通話先システム上で取得した各種情報、ルール情報を記述した知識情報とを評価することで、会話中に現れる重要なキーワードの抽出と、そのキーワードに関連する情報を決定する。このとき、自システム上の情報だけではなく、通話先システム上で取得した各種情報を用いることで、例えば、自分が「新宿」、通話相手が「8時」と言った場合に、両者の発話を合成し、新宿に8時に到着するための電車の時刻表を両者の端末上に表示することが可能となる。また、ユーザの会話パターン等からユーザの特徴を把握し、知識情報の更新を行う。例えば、会話中に特定のキーワードが多く現れる場合には、そのキーワードに関連のある情報が他の情報よりも優先して提供されるようになる。

情報表示機能は、コンテキスト評価機能で出力された情報や情報共有機能で取得した情報と、通話相手の画面上に表示されている情報とを画面上に表示する。

情報共有機能は、音声認識結果やナビゲーションシステムによって提示された提供情報を通話先システムと共有する。現状の音声認識システムでは、特定話者に対する認識率のほうが不特定話者に対する認識率と比較して高いため、各自の端末で音声認識を行い、音声認識結果を交換することで音声認識率の問題をある程度解決することができる。さらに、自システムでの音声認識結果の送信の可否を決定できるため、プライバシーの問題も同時に解決している。

各システムが持つ知識情報は必ずしも同じではないため、同一の入力情報に対して、異なる提供情報を出力する場合がある。このため、提供情報自身も情報共有機能経由で交換することで、互いに同一の情報を閲覧することが可能となる。これにより、あるユーザの画面上に表

示された情報を通話相手と共有でき、通話相手と同一の情報を閲覧しながら会話を行うことができる。

5. 現在の状況と今後の課題

以上述べた会話連動ナビゲーションシステムのプロトタイプを実装した。複数の会話シナリオに基づいて知識情報を構築し、動作の確認を実施した。その結果、提案システムによる会話に対する情報提供の実現性を確認することができた。

今後、一般的な会話に対する情報を提供するシステムを構築するために解決すべき課題を以下に示す。

- (1) 知識情報作成と拡張の自動化、知識の正確さの向上
- (2) 音声認識システムにおける誤認識への耐性向上
- (3) 会話中の話題の変化に追従した情報提供方式
- (4) 知識情報を大規模化した際の実時間処理の確保

6. おわりに

会話音声と背景情報(コンテキスト)に関連した情報提供を行う会話連動ナビゲーションシステムの提案を行った。また、プロトタイプの作成を実施し、プロトタイプを試用した結果から、本システムを一般的な会話に適用させるために必要となる課題の整理を行った。今後は、知識情報を大規模化し、一般的な会話に対する情報を提供するシステムを構築する予定である。

参考文献

- [1] 電子情報技術産業協会, “2003年度音声入出力方式に関する調査研究報告書 付録2 音声認識の製品動向”, http://tsc.jeita.or.jp/TSC/COMMS/4_IT/Speech/DATA/recog-H15.pdf
- [2] NTTコミュニケーションズ Vポータル, <http://www.ntt.com/v-portal/>
- [3] 大前 壮司 他, “実環境音声情報案内システムにおける発話感情理解についての検討”, 日本音響学会研究発表会講演論文集, vol.1, pp.205-206, 2004-9.
- [4] 後藤 真孝, “非言語情報を活用した音声インタフェース”, 情報処理学会研究報 2004-SLP-52-7(2004).
- [5] 城塚 音也 他, “エージェントアーキテクチャに基づいた音声対話支援システムの試作”, 情報処理学会研究報告 97-SLP-15-18 (1997).