

# 音声対話技術

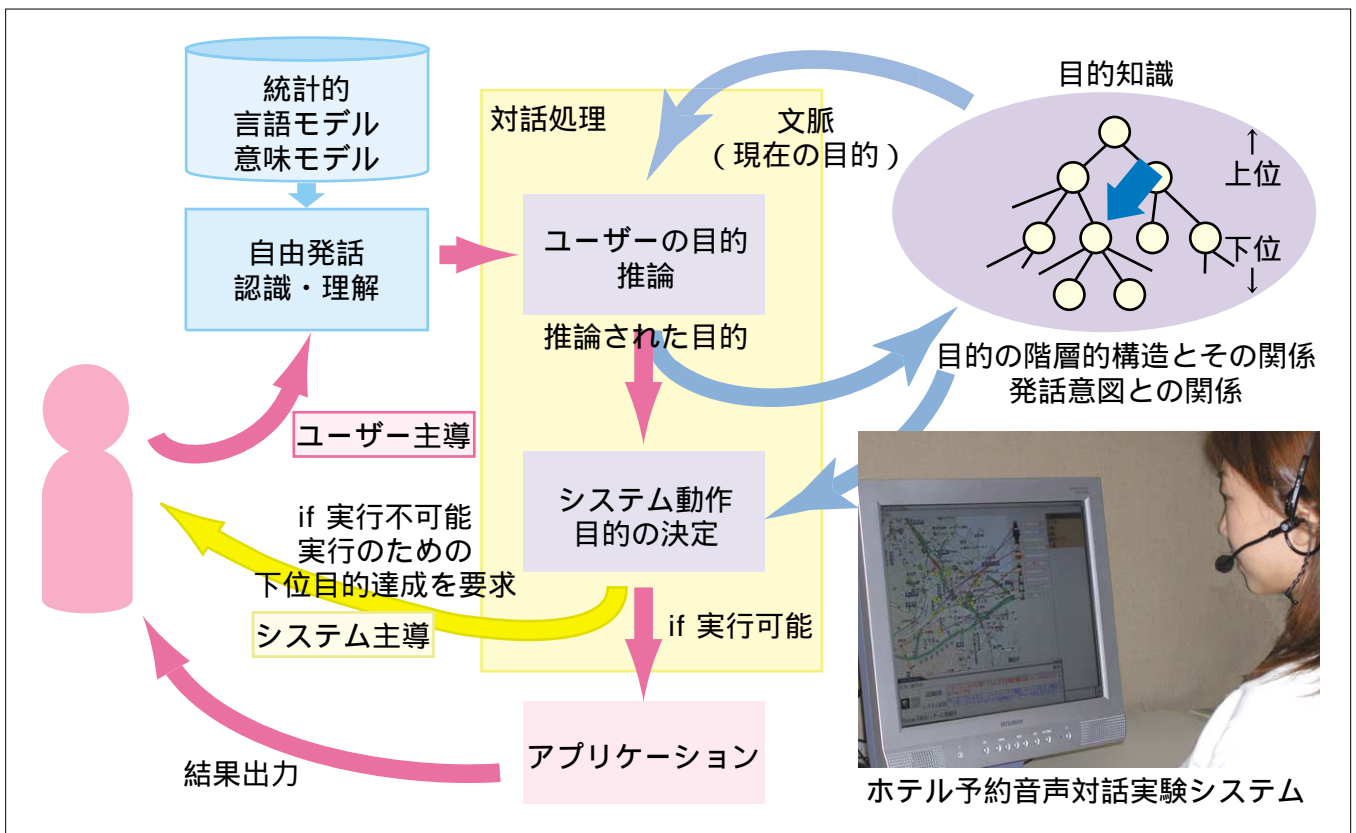
石川 泰\*  
 渡邊圭輔\*  
 谷垣宏一\*

## 要 旨

音声は人間にとって最も基本的な情報伝達手段であり、機械を利用する場合も、人間に話し掛けるように意志を伝え会話をしたいという要求は極めて自然なものである。この音声によるインタフェースを構築する上で重要な技術課題が対話処理である。対話処理とは、ユーザーとシステムとの間の入出力を管理し、ユーザーの目的を適切に達成させる対話の維持管理と、前後の発話やシステムの状況に依存したユーザーの発話意図を把握する文脈処理を行うものである。このうち、対話の維持管理では、システムが対話の主導権を持ち、ユーザーに決められた手順で発話を促す方法があるが、利便性が極めて低いものとなる。これに対し三菱電機は、基本的にユーザーに主導権を与え、かつ、目的達成を支援するとともに、文脈処理を実現する対話管

理方式を開発している。この方式では、個々の発話意図からユーザーの対話を通じての上位の目的を推論し、システムはその目的達成のための動作を行う。達成のために必要な下位の目的があれば、システムが対話の主導権をとり、下位目的を実行させることで、効率的に目的を遂行できる。さらに、目的の達成状態が文脈を表すため、文脈処理も同時に実現している。

本稿では、ホテル予約をタスク(対象業務)とする実験システムを例に、この方法を説明する。さらに、関連技術である対話エージェント“COLLAGEN”の紹介を交えて、音声対話技術の今後の課題と、対話を観点とする将来のマンマシンインタフェースについて展望する。



## 音声対話のブロック図とホテル予約実験システム

音声対話システムでは、自由に発声した会話文を理解するための統計的言語モデル・意味モデルが適用され、入力音声から発話意図とパラメータからなる発話意味が抽出される。対話処理部では、発話意図からユーザーの目的を推論し、それに応じたシステムの動作目的を決定する。システムが動作目的を達成できない場合、必要な下位の目的を、システムは、動作目的として、対話の主導権を得て対話を進める。