

デー05 対話を用いてユーザの発話内容を推論する機器操作システム

長田 誠也、土井 伸一、亀井 真一郎
NEC C&Cメディア研究所

はじめに

音声、言語によるコミュニケーションは人間にとって最も自然なものである。最近では、音声でパソコン上のアプリケーションを操作するシステムや、音声で操作可能なカーナビゲーションシステム等も実用化されてきている。

理想的には、単にあらかじめ決められたコマンドの音声発話を認識できるだけでなく、種々の言い回しや省略等の表現を解釈できるインタフェースが望ましい。しかしこの場合、ユーザの発話は一般的には曖昧でシステムが一意に理解できるとは限らない。そこでユーザと対話を行うことによって解釈を確定する機構が必要となる。ただし、システムからの応答が常にユーザに「Yes/No」を要求するインタフェースは自然なものとは言えない。そこで著者らは「確信度」というパラメータを用いて、適切なシステムからの応答文を生成する機構を導入した。

本稿では、著者らが試作した音声入力を解釈してテレビとビデオを操作対象とする機器操作システムの全体構成、およびシステムからの応答文と対話について、順に報告する。

システムの概要

本システムの構成を図1に示す。各モジュールの概要は以下の通りである。ユーザが「つけて」という曖昧性のある発話したときを例にとって説明する。

システムは「発話入力部」で「つけて」というユーザの発話を音声認識する。

「入力解析部」では「知識保持部」の知識を用いてユーザの発話を解釈し、その解釈の確からしさを

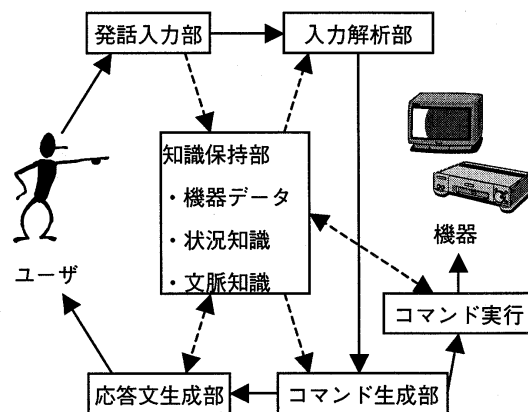


図1 システムの全体構成

表す「確信度」を求める。「つける」の場合「テレビの電源をつける」、「ビデオの電源をつける」等の解釈の曖昧性がありえるが、「知識保持部」の知識を用いて解釈を一意にする。例えば「テレビの電源をつける」と解釈したとする。

「コマンド生成部」では実際に「テレビの電源をつける」ためのコマンドを生成し、「応答文生成部」で「確信度」に応じた応答文の生成を行う。「確信度」が最も高い場合には機器操作を行った後に「テレビをつけました」と応答をし、「確信度」が低い場合には「テレビの電源をつけますか？」のようにユーザに質問をする。

システムからの応答文と対話

入力解析部では発話に曖昧性がある場合、その中から一番もっともらしい解釈を選択する。その際、その解釈の確からしさ（確信度）に応じて、システムからの応答を以下のように変化させる。

An electric appliances operating system that infers contents of user's utterances using dialog

Seiya OSADA, Shinichi DOI and Shin'ichiro KAMEI

NEC C&C Media Laboratories

〒216-8555 4-1-1 Miyazaki Miyamae-ku Kawasaki Japan

高	テレビをつけました。	事後報告
↑	テレビをつけます。	事前報告
確信度	テレビをつけますね。	確認
↓	テレビをつけますか？	Yes/No 質問
低	何をつけますか？	Wh 質問

確信度がもっとも高いときには、実操作を先にを行い、その後に「事後報告」を行うことにしている。ただしこの場合でも、直前の操作コマンドを文脈知識として記憶しているので、ユーザは簡単な発話で操作を取り消して元の状態に戻ることができるようになっている。

確信度が十分に高くないときには、実際に操作してよいかどうかをユーザに「確認」する。この応答に対してユーザが肯定も否定も行わない場合は、暗黙のうちに了解されたものとして操作するようにしている。これにより、ユーザにストレスを感じさせない自然な機器操作インタフェースを実現している。

さらに確信度が低いときに限っては、操作を行う前に実際に操作を行ってよいかどうかをユーザに「質問」する。

ここで、パソコン上での画面と対話例を以下に示す。図2は対話システムと仮想的なテレビとビデオである。実際のシステムでは仮想機器から実機器に

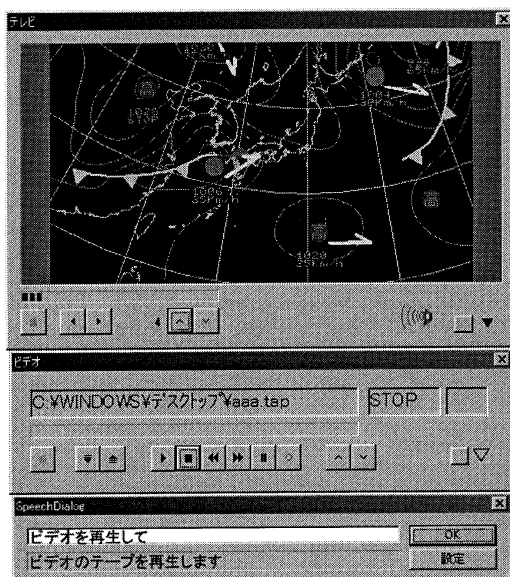


図2 対話システムと仮想機器

リモコンの信号を送ることで実機器を操作している。

(U：ユーザの発話、S：システムの応答)

・ U：「つけて」

S：「テレビの電源をつけますか」

※ユーザの入力に曖昧性があるので、システムからの応答文は質問形式をとる

U：「はい」

S：「テレビの電源をつけました」

※システムの質問形式の応答とユーザの返答から、システムはユーザの意図を推論して実行する

・ U：「NHK」

S：「テレビのチャンネルをNHKにしますね」

※文脈情報を用いてユーザの発話を解釈し、システムの応答文は「確認」

U：「違う、ビデオ」

S：「取り消します。ビデオのチャンネルをNHKにしますね」

※簡単な発話で誤った解釈の取り消し修正

おわりに

本稿では、筆者らが試作した音声による発話を理解して機器操作を行うシステムについて報告した。本システムは、ユーザの発話内容を理解した結果に対して「確信度」というパラメータを用いることによって適切な応答を返すところに特徴がある。この機構によりストレスを感じさせずにユーザの意図を解釈するための対話を促進することができる。

今後は、対象となる機器の数や機能、対応する語彙をさらに拡充するとともに、実際に使用しながらの評価を行って、ユーザにとってより自然な形の機器操作インタフェースの実現を目指す。

参考文献

長田 他：「自然言語を用いて家庭機器操作を行う対話システム」，信学技法 SP98-73，1998.

茶園 他：「ユーザの注目情報を利用したマルチモーダル・インタフェースの試作」，情処第 58 回全大 5E-05，1999.

土井 他：「知識に基づく音声言語理解システムとしての機器操作インタフェース」，言語処理学会第 5 回年次大会，1999，pp.377-380