

G-011

分散型対話管理による複数話者対話システムの構築

Development of a Dialogue System
with Distributed Dialogue Manager for Multi-speaker鈴木 智彦[†] 大下 義弘[†] 小林 恵太[‡] 関口 芳廣[§] 西崎 博光[§]

Tomohiko Suzuki Yoshihiro Oshita Keita Kobayashi Yoshihiro Sekiguchi Hiromitsu Nisizaki

1. はじめに

近年、複数の話者を対象とした対話管理や、ロボットに関する研究 [1][2] が進んできた。

筆者らは複数の話者との対話を可能にするため、話者ごとに独自の発話意図系列を用意し、話者間の意図の同期をとりながら対話を制御する分散型対話管理法を提案した [3]。本稿では分散型対話管理法を利用した会議室予約を支援するための対話システムを構築し、その評価を行なった。3人の話者との対話実験では約95%の発話意図特定精度を得ることができた。

2. 分散型対話管理法

分散型対話管理法の処理の流れを図1に示す。

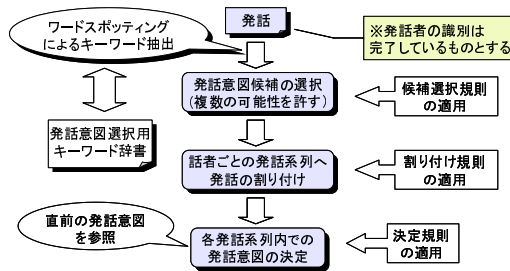


図1: 分散型対話管理の処理の流れ

分散型対話管理法では、まずキーワード辞書を基にワードスポッティングによるキーワード抽出を行ない、キーワードに対応した発話意図（複数個）を列挙していく。その中から発話意図候補選択規則を適用することで発話意図候補を決定する。次に、選択された発話意図候補を基に、話者ごとの発話意図系列へ発話を割り付ける。そして、各発話意図系列内において1つ前の発話意図（候補）と最新の割り付けられた発話の発話意図候補との組合せを考えることにより、発話意図の決定を行なう。なお本研究で使用する辞書、規則は人間同士の模擬対話を基にあらかじめ用意しておく。

3. 音声対話システム

分散型対話管理法による音声対話システムを作成する。音声対話システムの構成図を図2に示す。

音声対話システムは発話取得部、対話管理部、応答生成部の3つから構成される。発話取得部では音声認識ソフトを、応答生成部では音声合成ソフトを組み込んでいる。対話管理部では提案した分散型対話管理法を適用し、複数話者の対話を理解・管理している。

[†]山梨大学大学院医学工学総合教育部コンピュータ・メディア工学専攻

[‡]山梨大学工学部コンピュータ・メディア工学科

[§]山梨大学大学院医学工学総合教育部

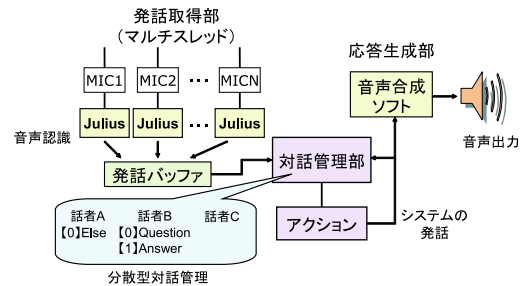


図2: 音声対話システムの構成図

3.1 タスク

タスクを「話者のスケジュール理解及び会議時間・部屋の決定」とする。話者が互いにスケジュールについての対話を行ない、システムはその対話の中から各話者のスケジュールを理解する。不足した情報や聞きとれなかった情報について質問を行なう。またシステムは会議に使用できる部屋、時間帯についての情報を保持しており、話者のスケジュールと照らし合わせ、最終的な結果として会議時間・部屋を決定する。

3.2 対話管理部

次に対話管理部の概要を図3に示す。

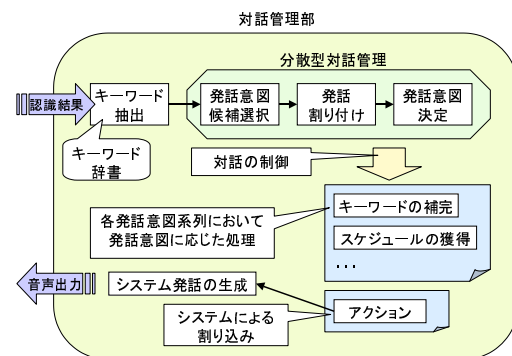


図3: 対話管理部の処理の流れ

対話管理部では、音声認識結果から抽出されたキーワードを基に発話意図候補を選択し、分散型対話管理を行なう。また、対話管理状況にあわせ、発話意図に関する情報の抽出やアクションを行なう。

3.2.1 キーワード抽出

ワードスポッティングによるキーワード抽出のための辞書は次の2つから構成される。

- 動詞辞書と名詞辞書
LOC (場所), ACT (活動) などの意味属性ごとに名詞を分類し, 動詞は格構造で意味を記述してある.
- キーワード辞書
発話意図候補の選択に関連するキーワードとキーワードの意味属性から構成される.

3.2.2 スケジュール

話者のスケジュールはスロットフィル構造をしており, slot とそれに対応する value によって表現される.

また, 話者が単語を省略して発話する場合もある. 特に, 話者が質問に対して答える場合にこのような現象が多い. そこで, 返答時に省略された単語を質問から補完し, 対話内容の理解に役立てることもしている.

3.2.3 アクション

システムは状況によって話者に対してアクションを起こす. このアクションは主に話者への質問発話となる.

このシステムではシステムの発話も一話者の発話として捉え, 話者の発話と同様の処理を行なう.

3.3 対話の制御

システムは以下の3つの Scene から成るスクリプトをもとに対話を行なう.

Scene1 では, 話者主導によるスケジュールに関する対話が行なわれる. システムは主に話を聞く立場にあるため, ランダムに対話に割り込むこと, 話者の発話に対して適切なあいづちをうつこと程度にしている.

Scene2 では, Scene1 で取得した内容に対してシステムがそれぞれの話者に質問・確認を行ない, 不足した情報を得る.

Scene3 では, 取得した情報よりシステムが会議時間を提案し, 話者の希望・スケジュールの変更等を受けつつ, システム主導の対話によりタスクを達成する.

4. 評価実験

構築した音声対話システムについて評価実験を行なった. 被験者数は成人男女10名である. このうち1対話につき任意の3名が参加し, 合計5対話の実験を実施した.

4.1 評価方法

次の5つの項目について評価した.

- WER: 音声認識時における単語誤り率.
- KER: キーワードの抽出誤り率.
- CER: 発話理解の誤り率. このシステムでは slot-value ペア取得精度に関係している.
- CER+: 省略されたキーワードの補完等を含めた対話理解度に関係した誤り率.
- 発話意図特定精度: 発話意図の推定誤りは, システムが間違ったスケジュールの取得や間違った質問をする原因に繋がる. そこで, このシステムの評価の一つとして, 役割別に発話意図特定精度を調査した.

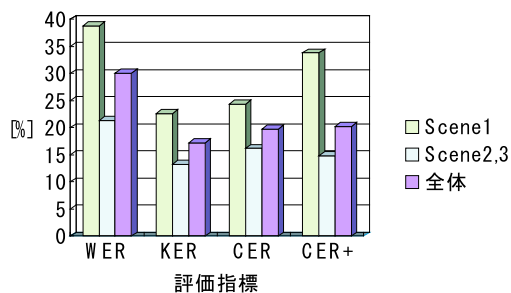


図4: シーン別誤り率

4.2 実験結果

各シーンにおける誤り率を図4に示す.

CER, CER+ともに全体で約20%程度と, ほぼ実用的に使用できる対話理解度を示している.

次にスケジュールの獲得状況を表1に示す. 話者が話した37スケジュール中, 35スケジュールを正しく認識できている.

表1: スケジュール獲得結果 (個)

スケジュール数	正解獲得数	間違い数	未取得数	否定数*
37	35	1	1	8

* 話者の否定で対話中に修正される.

次に発話意図特定精度を表2に示す. 結果より約95%の発話意図を正しく認識できている.

表2: 役割別の発話意図特定精度

		Scene1	Scene2,3	全体
発話者	正解率	69.2%	91.5%	83.9%
	正解精度	59.1%	100.0%	91.2%
聴者	正解率	86.7%	100.0%	97.9%
	正解精度	93.3%	100.0%	98.6%
観察者	正解率	93.3%	100.0%	98.6%
	正解精度	91.5%	100.0%	98.2%
全体の正解率		86.1%	98.7%	95.8%

5. おわりに

本研究では, 複数話者の対話を管理する分散型対話管理法を利用し, 音声対話システムの構築を行なった. その結果, 発話意図特定精度は約95%となり, 分散型対話管理法が複数話者を対象とした対話管理に有効であることが示された.

参考文献

- [1] 宮沢昇, 平沢純一, 中野幹生, 相川清明: 複数話者の発話が混在する発話系列を用いた対話理解, 情報処理学会, 音声言語処理研究会, No.2000-SLP-032-4, pp.19-24, 2000-07
- [2] 松坂要佐, 東條剛史, 小林哲則: グループ会話に参与する対話ロボットの構築, 電子情報通信学会論文誌 Vol.J84-D-II, No.6, pp.898-908, 2001
- [3] 大下義弘, 関口芳廣, 西崎博光: 複数話者対話システムのための分散型対話処理, 第4回情報科学技術フォーラム講演論文集, G-012, pp.289-290, 2005, 9