

# 会話エージェントにおけるマルチモーダル コミュニケーションのための対話管理プラットフォームの構築

山岡 雄治<sup>†</sup>中野 有紀子<sup>‡</sup>東京農工大学工学府情報工学専攻<sup>†</sup>成蹊大学理工学部情報科学科<sup>‡</sup>

## 1. はじめに

近年、人間の行動を計測するデバイスの発展により、自動計測されたユーザの行動データから得られた情報を、会話エージェントの振る舞い決定に利用することが可能になった。しかしながら、それらのデータを処理し対話管理を行う、つまりマルチモーダルな対話管理を行う手法については、研究者間で共有できるプラットフォームというものはなく、各システムに応じて、独自に作成しているのが現状である。そこで、本研究では、会話エージェントのように高度にマルチモーダル化された対話システムにおいて、大量に取得される人間の行動データを利用した対話状態更新機構を提案・実装する。

## 2. 会話の状態管理モデル

会話エージェントにおける、対話管理部の実装方法として、InformationState (以下 IS) [1]がある。IS には、現在の対話の目的、ターン保持者、発話の基盤化の状況、対話履歴による文脈等が保持され、IS 内の記述内容が更新されることにより、対話の進行が表現される。従って、時間  $t$  と  $t-1$  の IS を比較することにより、対話状態を比較することができる。この IS による会話モデルに準じた対話システムのプラットフォームの 1 つとして Midiki が公開されている。Midiki では、InformationState の変数を動的に格納することが可能であり、InformationState のカスタマイズを支援する機能が付加されている。一方、midiki は、テキストベースの対話システムを想定して作成されたツールであり、複数のモードからの入力を処理し、IS を更新するための機構については検討されていない。

## 3. マルチモーダル化に対応した IS の実現

Midiki の設計思想を踏襲しつつ、様々なモードの入力に適用しやすい対話管理機構を実現することを目的とし、IS 設定用の XML ファイルを定義することにより、IS の処理機構をカスタマイズする方法を提案する。本手法により、以下に示す表 1 の機能を実現する。

表 1 機能

機能	機能内容
IS で保持する情報のカスタマイズ	従来の midiki と同様、IS をシステムごとにカスタマイズできる機能をマルチモーダル化することにより、複数のモダリティ情報を自由に組み合わせたい IS を作成できる
モジュール間の処理変更	各モジュール間の処理フローを自由に設定できる
対話参加人数	エージェント、およびユーザの人数を設定可能

IS の定義ファイルの例を図 1 に示す。IS は、図 1 中で Cell と表記したセルを持つ。このセルには、名前をつけることが可能であり、また、セル内に定義された Speaker, Addressee といった type には変数をバインドすることも、さらにセルを埋め込むことも可能である。例えば、Cell1~4 の関係は、以下の図 2 のような階層構造を持つ。

```
< InfoState Version="0.0.0">
  <Cell ID="1" name="cVerbal">
    <variable varName="Speaker" type="obj"/>
    <variable varName="Addressee" type="obj"/>
  </Cell>
  <Cell ID="2" name="cNonVerbal">
    <variable varName="cUser1" type="cell"/>
  </Cell>
  <Cell ID="3" name="cUser1">
    <variable varName="Gaze" type="cell"/>
  </Cell>
  <Cell ID="4" name="Gaze">
    <variable varName="Time" type="obj"/>
    <variable varName="CorX" type="obj"/>
    <variable varName="CorY" type="obj"/>
  </Cell>
</DialogManager>
```

図 1 IS 定義 xml ファイルの例

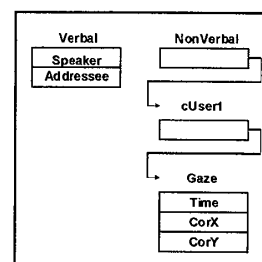


図 2 IS の階層構造

言語情報を保持するための verbal セルは, Speaker と Addressee という値を持つ. それに対し, 非言語情報を保持する NonVerbal セルは, cUserCell1 というセルを持ち, そのセルは, 視線情報を保持する Gaze セルを持つ. Gaze セルは Time (時間) と CorX (注視位置の x 座標), CorY (注視位置の y 座標) という 3 つの値を持つ.

#### 4. 提案するシステムアーキテクチャ

提案手法による IS を用いた対話管理プラットフォームとして, 図 3 に示す会話エージェントのための対話制御部のアーキテクチャを提案する. 以下では, ユーザの視線情報を利用したシステムを例にあげ, 本アーキテクチャの各モジュールについて説明する. 実装例のシステムでは, ユーザの視線推移情報からユーザの会話参加態度推定を行うシステム[2]を組み込み, 視線情報から推定されるユーザの会話参加態度に応じて会話エージェントの振る舞いを変化させる. このシステムでは, 120Hz で得られる視線情報と, 数秒に 1 回更新される発話情報の同期をとりながら IS を更新することにより, 粒度の異なる言語・非言語情報の一元的な管理を実現している.

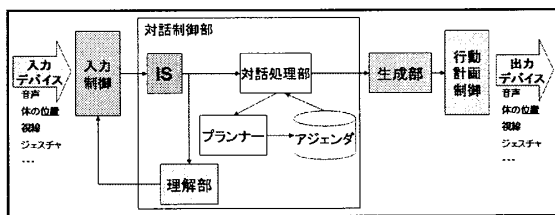


図3 提案するシステムアーキテクチャ

対話制御部は, IS の更新とエージェントの行動決定を行う. IS の更新にマイクやアイトラッカーから得られたデータを IS にて登録・更新するローデータ管理, 発話内容, 話し手, 聞き手等の発話情報の管理がある. 例えば, 視線データの場合は, ローデータ管理において, 視線位置と時間が更新される. 更新データは, 先に述べた ISxml ファイルに定義可能である.

##### (a) 理解部の情報管理

IS が視線のローデータによって更新されると, 次に, ローデータから視線推移をクラスタリングし, 設定した閾値によりユーザの会話参加態度を推定し, その結果を用いて再び IS を更新する.

##### (b) 会話参加態度を考慮した対話制御

対話処理部は, 更新された IS に基づいて, システム発話終了時に, 次のエージェント動作を決定する. その際, 理解部からのユーザの会話参加態度低下情報もしくは, ユーザの発話による IS の更新が, 行われることによって, 対話処理部に割り込みが発生する. 例えば, ユーザの会話参加度低

下が検出された場合, これが割り込みとなり, エージェントは「ここまでで質問はありますか?」といった, ユーザに興味を引かせる質問を行う.

##### (c) プランナーとアジェンダ

対話のゴールとエージェントの次動作の決定を行う.

#### 5. エージェントとの対話例

以上の機能を実装した会話エージェントとの対話の流れを図4に示す.

[G1:S] いらっしゃいませ. …	
[D1:S] 左上に …D904i です. … (ユーザの質問)	
[DQ1:H] D904i の価格は? [DA1:S] D904i の価格は新規契約で 36500 円です. [Q1:S] 説明を続けてもよろしいでしょうか? [QA1:H] はい. (再び D904i の説明) …	
[D27:S] ワンセグ対応していないのが少し不満です. (次の携帯説明へ)	
[L1:S] 右中央に …SIMPUREL2 です. … (ユーザが飽きて、逸脱した視線行動が表出)	
[Q2:S] 説明を続けてもよろしいでしょうか? [QA2:H] 次の携帯の説明をお願いします. (話を次の携帯の話へ)	
[K1:S] 次に, 左下にございますのが, 安心機能満載のキッズケータイ SA800i です. …	
[K10:S] カラーリングも豊富で, お子様も喜ぶこと間違いありません. [E1:S] またのお越しをお待ちしています.	

図4 エージェントとの対話例

#### 6. おわりに

本稿では, 提案するアーキテクチャをユーザの視線推移情報からユーザの会話参加態度推定[2]を行うシステムに組み込み, 実装を行った. 今後, WizardOz によって得られた評価と本システムの評価結果を比較することにより, 本アーキテクチャの有効性を確かめる. また, 本アーキテクチャを他のエージェントシステムに組み込むことにより, 汎用性の評価も行う予定である.

#### 7. 参考文献

- [1] David Traum et al. , A VIRTUAL HUMAN DIALOGUE MODEL FOR NON-TERM INTERACTION in Recent Trends in Discourse and Dialogue, 2008
- [2] 石井亮, 中野有紀子, ユーザの注視行動に基づく会話参加態度の推定—会話エージェントにおける適応的会話制御に向けて, Estimate Gaze 2008 年 12 月 Vol. 49 No. 12, 3835-3846