

動的環境情報を利用したマルチモーダル対話システムの構築と 適応応答文生成法の検討*

吉田 昌弘[†] 大隈 祐治[†] 桂田 浩一[†] 新田 恒雄[†]

[†]豊橋技術科学大学 大学院工学研究科 知識情報工学専攻

1. はじめに

近年、携帯電話や PDA・タブレット PC 等ユーザ端末の多様化に伴い、Web アプリケーションのマルチモーダル対話 (MMI) への応用が活発に議論されている。MMI では様々な入出力手段が用いられるが、その中でも、音声は数多くのシステムで利用されており、中核技術の一つとよいてよい。従来の音声対話システムでは、例えば VoiceXML [1]ベースのシステムのように、予め固定した対話シナリオを用意し、それに従った対話進行を行うのが一般的である。しかしこのようなシステムでは、どのようなユーザがどのような環境で発話しても、同じ応答を返すことしかできないため、対話が単調になってしまうという問題が生じる。本研究では、ユーザや環境に適応した応答文を生成することで、より自然な対話を実現する MMI システムの開発を目指す。

以上に述べたシステムを開発する上で、端末情報や、環境情報、ユーザプロフィール等の情報 (これらを動的環境情報と呼ぶ) の利用は必要不可欠である。Web 技術の標準化団体である W3C では、動的環境情報を web ベースシステムで利用するための枠組みとして、Delivery Context Interfaces (DCI) [2]の検討を行っている。本稿では、DCI を参考にして設計した動的環境情報管理部を、我々がこれまでに開発してきた MMI システム [3]に導入する。以下では、システムアーキテクチャを説明した後、ユーザや環境に適応した応答文を生成する手法を検討したので、報告する。

2. MMI システムの構成

我々がこれまで開発してきた MMI システムの構成を図 1 の破線部に示す。MMI システムは、入出力を管理するフロントエンド部、複数の入力モダリティを統合する入力統合部、ユーザへの出力を生成する出力生成部、シナリオに沿って対話を制御する対話制御部から成る。入力統合部・出力生成部は、それぞれに用意された統合ルール・生成ルールを用いて統合・生成を行う。また、対話シナリオはこれまで我々が検討してきた XISL2.0 (eXtensible Interaction Scenario Language) [4]によって記述し、対話制御部の要求に従ってドキュメントサーバからダウンロードする。

3. 動的環境情報管理部の導入

MMI システムにおいて動的環境情報を管理・利用するために、図 1 下部に示す動的環境情報管理部を追加した。動的環境情報管理部では、フロントエンド部から取得する端末情報、環境情報、ユーザプロフィールデータベースからダウンロードするユーザプロフィール等をアプリケーション毎に管理する。

動的環境情報は図 2 に示すように階層的に管理され、ルートから端末情報、ユーザプロフィールという様にツリーが分化している。ツリーを構成する各ノードをプロパティと呼び、子プロパティもその機能や性質ごとに分化してい

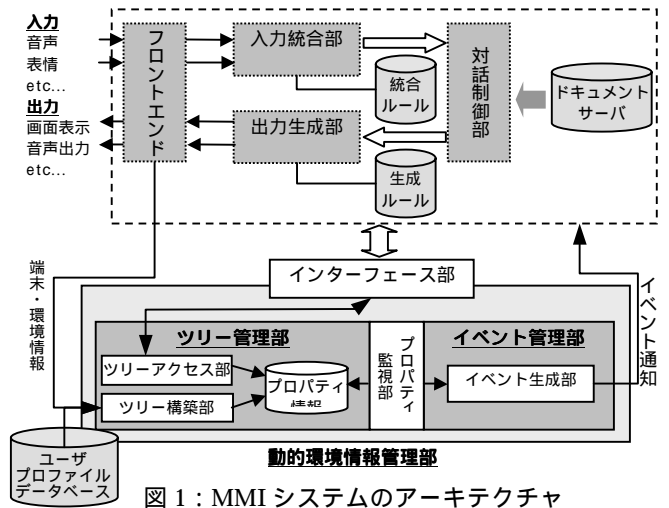


図 1: MMI システムのアーキテクチャ

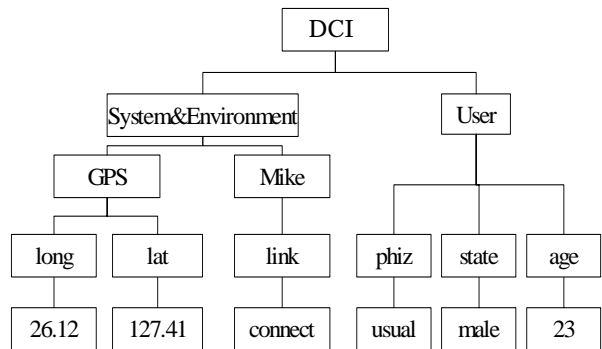


図 2: 動的環境情報の木構造の例

る。例えば、端末情報 (System&Environment) 下の GPS の子プロパティは緯度 (lat) と経度 (long) といったように構成されている。

動的環境情報管理部では、例えばマイクが ON から OFF に切り替わるというような動的な変化が起きた場合、フロントエンド部から送信された情報と、ツリー内のプロパティを比較して値を更新する。この時、同時にプロパティの変化を他のモジュールに通知するためにイベントを発生させる。

他のモジュールから動的環境情報を参照する場合は、DOM (Document Object Model) [5]を用いてプロパティツリーのノードを辿る方法の他に、プログラミング言語の関数のように、予め定義した予約語を指定することによって情報を取得する方法を提供している。前者の場合はツリーの構成を熟知している必要があるが、後者ではツリーの構成を気にする必要はないため情報の参照は比較的容易となる。

4. 応答文生成法の検討

本研究では、動的環境情報を利用してユーザに合わせた応答文を生成することで自然な対話を実現する。ここでは、そのための要素技術の一つとして、発話内行為から応答文を生成する方法を検討する。

*Implementation of a dynamic environmental-information manager and a proposal of an automatic response generation method in MMI systems.

A.Yoshida[†], Y.Okuma[†], K.Katsurada[†], T.Nitta[†]

[†]Graduate School of Engineering, Toyohashi Univ.of Tech.

4.1. 発話内行為

日常的な人間同士の対話では、発話がどのような意図に基づいて行われたかが重要となる。こうした発話の意図を取り扱う枠組みとして、オースティンやサルらは言語行為論[6][7]において、発話内行為を規定した。発話内行為とは、発話によってどのような行為が遂行されたのかを示すものである。例えば、「花を買ってあげるよ」という発話には、話者が聞き手に花を買うことを“約束する”という行為が含まれる。発話内行為の利用により、具体的な文としてではなく、抽象的な形で発話意図を記述することが可能になる。

しかし、発話内行為の種類は多数であるため、全てを対話システム内で準備することは極めて困難である。また、対話システムでは利用しないであろう発話内行為も多い。これらの理由から、本研究では対話システムにおける利用を前提として、発話内行為の種類を限定した。今回は、オンラインショッピングシステムでの利用を考え、12種類の発話内行為を選択した。例えば、ユーザがシステムを訪れたときや、システムを終了させる際の挨拶には“welcome”や“thank”，エラー時の警告には“warn”，説明時には“tell”を利用する。本研究で利用する発話内行為は文献[8]に挙げられている命名動詞から選択した。本研究で利用する発話内行為と場面を表1に示す。

4.2. 発話内行為を用いた応答文生成

応答文の生成には、発話の意図を示す発話内行為の他に、発話する内容が必要となる。本研究では、発話内容を発話行為に対する格情報として記述する。格情報のうち、述語のカテゴリ指定には情報処理推進機構 IPA で開発された計算機用日本語辞書 IPAL[9]の動詞辞書に記述されている分類名称を利用する。分類名称の指定により、ユーザや環境に応じて動詞を変化させることが可能となる。図3に本手法における発話内容の記述と生成される文の例を示す。

本手法における応答文生成の流れを図4に示す。対話シナリオ内には、発話内行為と文の格情報を記述する。応答文を生成する際には、まず発話内行為に依存した項目が記述されているかをチェックする。これは、例えば図3(b)(c)における“target”である。これは、文のどの部分を問うかを示す項目であり、この場合では文自体を問うことを示す。次に、発話内行為、格情報から該当する定型文があるかどうかを判断する。例えば、図3(a)に示すように歓迎を表すときには「いらっしゃいませ」や「ようこそ」、システム終了時には「ありがとうございました」といった定型文を用いる。次に、述語が記述されているかを判断し、記述されていた場合、IPAL から適切な動詞を取得する。例えば、図3(c)では、述語として「購買」が指定されているが、これに該当する動詞は「買う」や「購入する」といったものがある。こうして取得した動詞、格情報を発話内行為に合わせて整形する。この際、IPAL や EDR 日本語動詞共起パターン副辞書[10]等を利用し適切な助詞を補完することで文を生成する。

(a)	発話内行為: welcome ⇒ いらっしゃいませ
(b)	発話内行為: ask 発話内行為依存項目: target=sentence 目的語: りんご 修飾語: 4個 ⇒ 4個のりんごですね?
(c)	発話内行為: ask 発話内行為依存項目: target=sentence 述語: 購買 目的語: りんご 修飾語: 4個 ⇒ 4個のりんごを購入しますか?

図3: 記述内容と生成される文の例

表1: システムの場面と発話内行為

使用する場面の例	発話内行為
入力の促し	request
入力の承認	acknowledge
	engage
入力の否認	reject
挨拶	welcome
	thank
	greet
在庫切れ等に対する謝罪	apologize
エラー等の警告	warn
説明	tell
	negate
入力の確認等の質問	ask

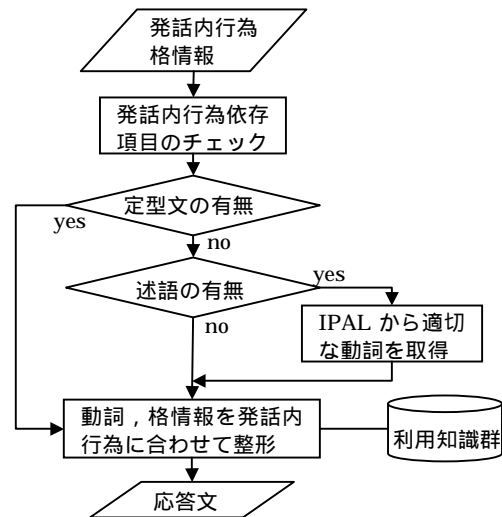


図4: 検討手法のフローチャート

5. おわりに

動的環境情報を MMI システムで利用するため、動的環境情報管理部を導入した。また、動的環境情報に基づく応答文生成のために発話内行為を利用する手法について検討した。これらの技術により、ユーザや環境に適応した応答文の生成が可能になる。

今後は、応答文をユーザや環境に合わせてスタイリングする方法を検討するとともに、システムの実装を進めていきたい。

参考文献

- [1] <http://www.w3.org/TR/voicexml21/>
- [2] <http://www.w3.org/TR/DPF/>
- [3] 桂田浩一, 他: “多様な端末からのアクセスを可能にする MMI アーキテクチャ”, 情処研報 2002-SLP-40, pp.51-56(2002)
- [4] 青木一峰, 他: “異なる端末環境から利用可能なMMIアプリケーション開発における記述負担の軽減”, 信学技報, SP2004-126, pp.25-30, (2004-12).
- [5] <http://www.w3.org/DOM/>
- [6] Austin, J.L: “How to do Things with Words”, Oxford University Press, (1962)
- [7] Searle, John: “Speech Acts: an essay in the Philosophy of Language”, Cambridge, (1969)
- [8] ダニエル・ヴァンダーヴェーケン: “意味と発話行為”, ひつじ書房, (1997).
- [9] <http://www.ipa.go.jp/>
- [10] http://www2.nict.go.jp/kk/e416/EDR/J_index.html