

ユーザ環境情報を利用した対話が可能なマルチモーダル対話システムの検討*

大隈 祐治† 青木 一峰† 桂田 浩一† 山田 博文‡ 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
 ‡豊橋技術科学大学 マルチメディアセンター

1. はじめに

近年、携帯電話や PDA・タブレット PC 等のユーザ端末多様化に伴い、Web アプリケーションへ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。さらに、最近では、端末情報、環境情報、ユーザプロフィール (年齢、性別、嗜好、...) (以上をユーザ環境情報と呼ぶ) を利用した MMI の実現が期待されている。web 技術の標準化団体である W3C においても、ユーザ環境情報を含む動的に変化する様々な情報を利用するためのフレームワークとして Dynamic Properties Framework (DPF) [1] の検討が始まっている。

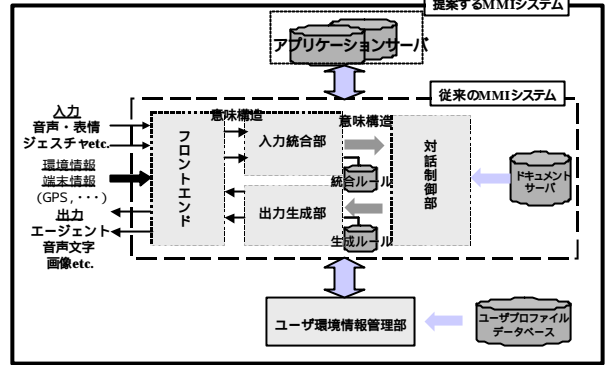
本稿では、我々がこれまで開発してきた MMI システム [2] に、DPF の考えを取り入れて設計したユーザ環境情報管理部を導入したので報告する。以下、我々が開発してきた MMI システムのアーキテクチャを説明した後、ユーザ環境情報管理部の概要および他のモジュールとの関係について、利用例を交えて述べる。

2. MMI システムのアーキテクチャ

図 1 (a) に MMI システムのアーキテクチャを示す。このアーキテクチャは図 1 (b) に示す U-M-A 三階層モデル [3] に基づいて設計した。U-M-A 三階層モデルは、MMI システムの基本構成を与えるもので、Uni-modal/Uni-medium 層, Multi-modal/Multi-media 層, A-modal (モダリティに依存しない) 層の三階層から成る。我々はこれら三階層をそれぞれ MMI システムのコンポーネントとして、フロントエンド部 (入出力を管理する、同時に端末情報・環境情報を収集する)、入力統合部/出力生成部 (複数の入力モダリティを統合する/ユーザへの出力を生成する)、対話制御部 (シナリオに沿って対話を制御する) として実装した。入力統合および出力生成はそれぞれ入力統合ルール、出力生成ルールの記述に従って実行される。また、対話シナリオはドキュメントサーバに格納され、対話制御部の要求に従ってダウンロードされる。

以上に述べたシステムに、今回、アプリケーションに応じて外部モジュールと通信できるようにした。また、ユーザ環境情報管理部とユーザプロフィールデータベースを導入してユーザ情報の管理を行えるようにした。

(a) MMIシステムアーキテクチャ



(b) U-M-A三階層モデル

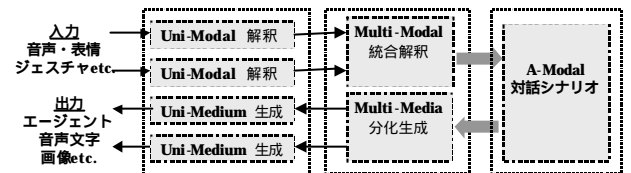


図 1. MMI システムアーキテクチャと U-M-A 三階層モデル

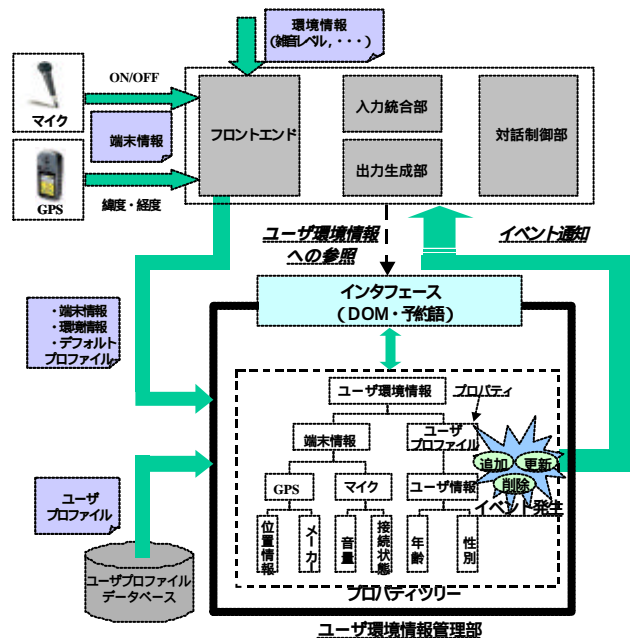


図 2. ユーザ環境情報管理部と他モジュールとの関係

3. ユーザ環境情報管理部

図 2 にユーザ環境情報管理部と他モジュールとの関係を示す。ユーザ環境情報管理部は、フロントエンドから送

*A proposal of an MMI system architecture allowing the use of user environment information

Y.Okuma†, K. Aoki†, K.Kat surada†, H.Yamada‡, T.Nitta†

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

‡ Multimedia Center, Toyohashi Univ. of Tech.

られる様々な情報(端末情報,環境情報,デフォルトプロファイル)の他,ユーザプロファイルデータベースからダウンロードしたアプリケーション毎のユーザ情報を管理する。以下にユーザ環境情報管理部の概要を示す。

3.1 ユーザ環境情報管理部のデータ管理方法

ユーザ環境情報は,階層的に管理され,図2に示すように,ルートから端末情報やユーザプロファイルのよう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, GPS やマイクなど子プロパティも機能や性質ごとに分化している。例えば GPS の子プロパティは位置情報とメーカー名,マイクの子プロパティは音量と接続状態,ユーザ情報の子プロパティは年齢と性別のように構成されている。

環境情報管理部は,まず,予め定義したツリー構造(デフォルトツリー)に従い,プロパティの値が空のツリー構造を構築する。その後,プロパティの値を埋めるために,アプリケーション開始時に,以下の内容を実行する。

ユーザプロファイルデータベースから,ユーザ情報(名前,年齢,性別,ユーザ嗜好など)を取得する。

フロントエンドからデフォルトプロファイル(当該端末の端末機能や環境情報)を取得する。

3.2 他モジュールへのイベント通知機能

マイクが ON から OFF に切り替わるといった動的な変化が起きた場合,ユーザ環境情報管理部は,フロントエンド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。イベント通知は全てのモジュールに対して行われ,特定のイベントの待ち受けやイベントの利用は各モジュールに委ねられる。

3.3 他モジュールへのインタフェースの提供

他のモジュール,または対話シナリオからユーザ環境情報を参照することを可能にするためインタフェースを二つ設けた。一つはツリーのノードを辿っていく DOM (Document Object Model) [4]を用いる方法である。これには「GPS プロパティの子プロパティの中から,アクティブなプロパティの値を取得する」のような記述を必要とするため,ツリー構造の熟知が必要である。もう一つは,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(=プロパティ名。例えば GPS:vender)を指定するインタフェースを用いる方法である。この場合は,予約語に対応するプロパティの値が自動的に返される。DOM のように複雑なツリー構造を指定する必要が無いため,ツリー構造を熟知せずに利用できる利点がある。

4. MMIシステムにおけるユーザ環境情報の利用

4.1 フロントエンドでの利用

ユーザ環境情報を用いることで,ユーザからの入力時やユーザへの出力時に,適切なインタフェースを提供することが可能になる。例えば,「周囲の騒音がある一定値を超

えていた場合,音声認識の音響モデルを雑音に頑健な音響モデルに切り替える」というように,環境に対応して音声認識モジュールを構成し,性能を向上させることができる。また,フォントの大きさや音量の調節,画面表示サイズなど,ユーザの嗜好や端末性能に合致した出力を実現することもできる。

4.2 入力統合部での利用

入力統合ルールでユーザ環境情報を参照できると,対話に必要な値を取得することが可能になる。GPS 情報に基づいて列車の予約をするコンテンツを例に取ると,ユーザが「ここから東京まで」と発話した際に「ここ」という発話を最依りの駅と捉えたいことがある。このとき,ユーザ環境情報管理部から GPS の情報を取得し,地図データベースモジュールへ GPS からの位置情報と建築物の種類を渡すため,以下のような記述を入力統合ルールに記述すれば,「最依りの駅」を埋めることができる。

```
<駅名>Submit.MAP(SE.GetGPS,"station")</駅名>
```

4.3 出力生成部での利用

出力生成ルールに,ユーザ環境情報を参照する記述ができると,例えば画面表示デバイスが無いときは,出力を画面から音声に変更するといった切り替えが可能になる。また,ユーザが低年齢の場合には,説明口調から話し言葉口調へとスタイルを変更するというように,ユーザ環境情報に応じた提示が可能になる。

4.4 対話制御部での利用

対話シナリオでユーザ環境情報からのイベントを待ち受ける記述や,ユーザ環境情報管理部が管理している情報を参照する記述が可能になると,例えば,ユーザの表情変化をイベントとして待ち受ける記述を加えることにより,対話戦略の切り替えや対話シナリオの変更が実現可能になる結果,より豊かな対話が可能になると考えられる。

5. おわりに

ユーザ環境情報を MMI システムで利用するため,ユーザ環境情報管理部の導入を検討した。この導入により,フロントエンドでのユーザインタフェースの変更から対話制御部での対話シナリオの変更まで,環境変化に対応した柔軟な対話構築が可能になる。

今後,フロントエンドとユーザ環境情報管理部間の通信プロトコル仕様策定すると共に,有効性を示す実行システムを開発する予定である。

文献

- [1] <http://www.w3.org/TR/DPF/>
- [2] 桂田浩一,他,“多用途な端末からのアクセスを可能にする MMI アーキテクチャ”,情報処理学会研究報告書 2002-SLP-40, pp.51-56(2002)
- [3] 新田恒雄,“マルチモーダル対話の深化と記述言語の今後”,情報処理学会研究報告,2004-SLP-50, pp.15-22, Feb.2004.
- [4] <http://www.w3.org/DOM/>