

多様な端末からのアクセスが可能な OLS システムの実装

桂田浩一* 小林剛典* 中島将宏* 中嶋真里子** 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 知識情報工学課程

***豊橋技術科学大学 マルチメディアセンター

〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1 - 1

Email: {katurada, kobayashi, nakajima, nakashima, yamada, nitta}@vox.tutkie.tut.ac.jp

あらまし： 筆者らはこれまでマルチモーダル対話 (MMI) 言語 XISL (eXtensible Interaction Scenario Language) を検討してきた。XISL はモダリティの拡張性が高いという特徴を持つため、多様な端末を用いた MMI システムでの利用が期待できる。今回、XISL のこうした特徴を実証するために、PC、PDA、電話を端末とする 3 種類の MMI システムを構築し、これら端末上で動作するオンラインショッピング (OLS) アプリケーションを試作した。PC を端末とするシステムは駅やコンビニエンスストアに設置される情報キオスクを、PDA と電話を端末とするシステムはモバイルでの利用を想定して導入した。本稿では各 MMI システムの機能を示すとともに、それぞれのシステム向けに試作した OLS アプリケーションを紹介する。

キーワード： マルチモーダル対話システム, XISL, OLS, 多様な端末

Implementation of OLS Systems Accessible from Heterogeneous Terminals

Kouichi KATSURADA*, Takanori Kobayashi*, Masahiro Nakajima*,
Mariko Nakashima**, Hirobumi YAMADA*** and Tsuneo NITTA*

*Graduate School of Engineering, Toyohashi Univ. of Technology

**Faculty of Engineering, Toyohashi Univ. of Technology

***Multimedia Center, Toyohashi Univ. of Technology

1-1 Hibarigaoka, Tempaku-cho, Toyohashi 441-8580, JAPAN

Email: {katurada, kobayashi, nakajima, nakashima, yamada, nitta}@vox.tutkie.tut.ac.jp

Abstract : We have investigated a basic MMI system architecture, and have specified an MMI description language XISL (eXtensible Interaction Scenario Language) for providing seamless web services. In this paper we provide MMI system implementation on three types of terminals: a PC, a telephone server, and a PDA. The PC models an information kiosk in a convenience store. In the proposed MMI system architecture, MMI scenarios written in XISL are interpreted, then interactions are executed according to the scenarios. We show that our approach enables a system designer to implement the same application in web services on heterogeneous terminals by showing a sample application of online shopping system.

Key words : Multimodal interaction system, XISL, OLS, Heterogeneous terminals

1. はじめに

近年、多様なモダリティを用いて Web にアクセスするための基盤技術が整備されつつある。また W3C においても、昨年 2 月にマルチモダルインタラクション (MMI) ワーキンググループ[1]が結成され、活発な議論が始まっている。同時に、その他機関でも独自の MMI 記述言語が検討されている[2][3][4]。

我々は 2 年前から MMI 記述言語 XISL [5][6][7] (eXtensible Interaction Scenario Language) と MMI システムアーキテクチャ [8]を検討してきた。XISL は入出力モダリティ記述に自由度を持つため、端末の拡張性が高いという特徴を持つ。また MMI システムは、端末に依存したモジュールと非依存のモジュールを明確に分離しているため、新規端末の導入が容易である。これらの特徴から、我々が検討してきた言語およびシステムを利用することで、多様な端末を利用したシームレスな Web サービスを実現することが期待できる。

本報告では、XISL および MMI システムのこれら特徴を実証するため、PC、PDA、電話を端末とする 3 種類の MMI システムを実装し、オンラインショッピング (OLS) のアプリケーションを作成した結果を述べる。各 MMI システムで用いる端末は性能に大きな開きがあるため、実装形態が大きく異なる。例えば各システムは画面の有無や大きさなど、ユーザインタフェース (UI) が異なるため、OLS アプリケーションもシステムの UI に合わせて適切な形を取る必要がある。本文では各 MMI システム向けに作成した OLS アプリケーションのそれぞれの特徴について説明する。

以下、まず XISL および MMI システムのアーキテクチャを簡単に紹介した後、各端末を用いた MMI システムの詳細、OLS アプリケーションの特徴の順に説明する。

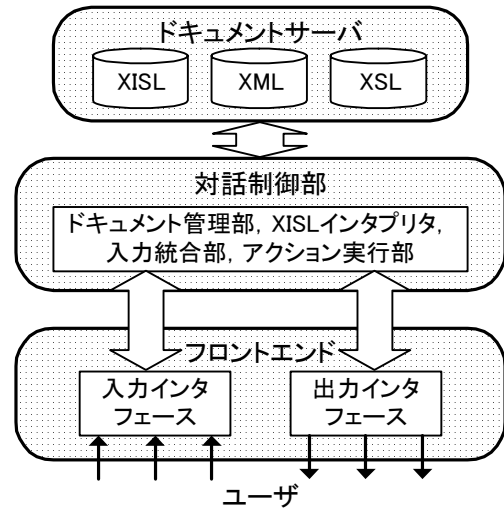


図 1. MMI システムのアーキテクチャ

2. XISL の概要と MMI システムアーキテクチャ

2.1. MMI 記述言語 XISL

XISL は XML 仕様に基づくマークアップ言語で、以下の条件を満たすよう策定した。

- 1) モダリティ記述の自由度が高いこと。
- 2) 多様なモダリティを組み合わせた入出力が記述可能なこと。
- 3) 対話記述言語として必要な機能を備えること。

まず 1) を満たすために、モダリティ依存の記述を XISL 仕様を含めず、端末の開発者が自由に規定できるようにした。続いて 2) を満たすために複数モダリティを用いた同時 / 逐次入出力、および択一入力を記述可能にした。さらに 3) を満たすために、対話の遷移、条件分岐、およびその他のアクションを記述可能にした。これらの特徴により、端末毎にモダリティ記述の仕様を定めさえすれば、多様な端末を用いた MMI システムで XISL を利用することが可能になる。

2.2. MMI システムのアーキテクチャ

端末の拡張性を高めるためには、XISL のモダリティ記述の自由度が高いことに加えて、多様な端末の実装を可能にする MMI アーキテク

チャを提供しなければならない。我々は MMI システムを構成するモジュールのうち、端末に依存するモジュールと端末に非依存のモジュールを明確に分離することにより、これを実現した。図 1 に MMI システムのアーキテクチャを示す。

フロントエンドは XISL 記述のうち端末依存の部分を解釈するモジュールである。このモジュールは、MMI システムの種類毎に設計・実装する必要がある。一方、対話制御部は XISL 記述のうち、端末非依存な部分のみを解釈・実行すればよい。ため、端末の異なる MMI システム間で共通に利用できる。したがって、システム開発者は新たな端末を導入する際には、フロントエンドのみを実装すればよい。このことは、新規端末の導入コストを低く抑えることに繋がる。

3. 多様な端末を用いた MMI システム

XISL および MMI システムアーキテクチャの有用性を実証するために、PC、PDA、および電話を端末とする 3 種類の MMI システムを構築した。以下、それぞれのシステムの簡単な構成と利用可能なモダリティを示す。

3.1. PC を端末とする MMI システム

PC を端末とする MMI システムは、駅やコンビニエンスストア等に設置される情報キオスクを想定して開発した。今回開発したシステムは、現在の PC の性能上、図 2 に示すようにフロントエンドと対話制御部を別々の PC 上に実装した。将来の PC 性能向上により、この部分は一台の PC 上に実装可能となろう。システムで利用できる入出力モダリティは以下の通りである。

入力モダリティ

- 音声入力 (SAPI 対応の認識エンジン + キーワードスポッティング[9])
- タッチ入力 (クリック, ダブルクリック)
- キー入力

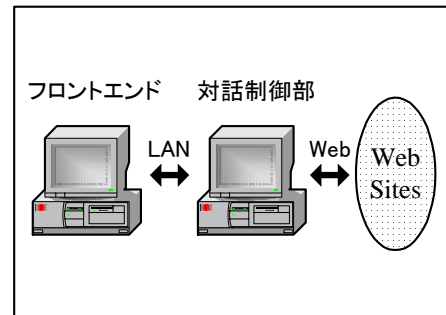


図 2 . PC を端末とする MMI システム

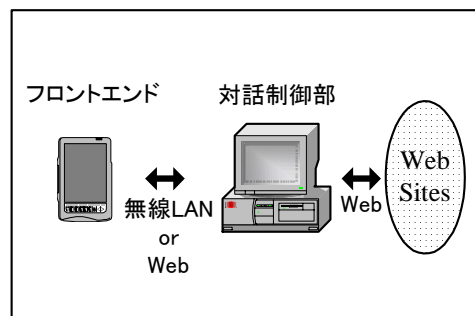


図 3 . PDA を端末とする MMI システム

出力モダリティ

- 合成音声出力 (TTS ; 東芝)
- 擬人化エージェント出力 (吹き出し, 発話, アクション等 ; MS エージェント)
- Web ブラウザ (表示, 更新, サイズ変更, 移動等)
- 動画再生

3.2. PDA を端末とする MMI システム

PDA は PC と比べて処理能力が劣るため、図 3 に示すように、PDA 上にはフロントエンドのみを実装した。対話制御部は別途 PC 上に実装し、無線 LAN もしくは Web を通じて PDA と通信を行う。利用可能な入出力モダリティは以下の通りである。

入力モダリティ

- 音声入力 (孤立単語音声認識 ; 東芝)
- ペンタッチ入力 (タップ, ソフトキーボード)
- カーソルキー入力

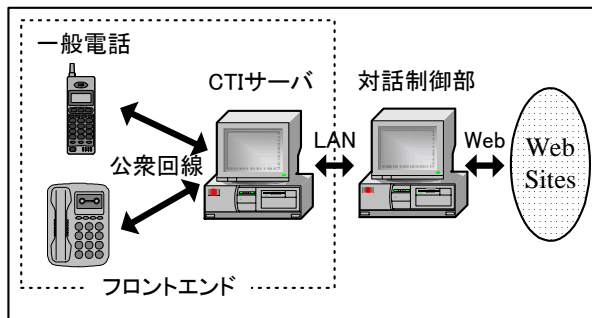


図 4 . 電話を端末とする MMI システム

出力モダリティ

- オーディオ出力
- Web ブラウザ (表示, 更新)

3.3. 電話を端末とする MMI システム

電話を端末とするシステムは,ボイスポータルサイトのような利用を想定して導入した.図 4 に示すように 2 台の PC 上に各々 CTI サーバと対話制御部を実装した.ユーザが固定電話もしくは携帯電話を用いて CTI サーバにアクセスすると,サーバが対話制御部と通信を行い,適切な XISL 文書を Web サイトからダウンロードする. CTI サーバは音声認識,音声合成といったフロントエンドの役割の大部分を果たす.利用可能な入出力モダリティは以下の通りである.

入力モダリティ

- 音声入力 (KDDI Speech Seeker + キーワードスポッティング[9])
- DTMF 入力

出力モダリティ

- 合成音声出力 (TTS ; 東芝)
- トーン出力

4. OLS アプリケーション

4.1. OLS アプリケーションの概要

前節で述べた各 MMI システム上で動作するオンラインショッピング(OLS)アプリケーションを開発した. OLS アプリケーションは,商品購入のための XISL 対話シナリオ,画面表

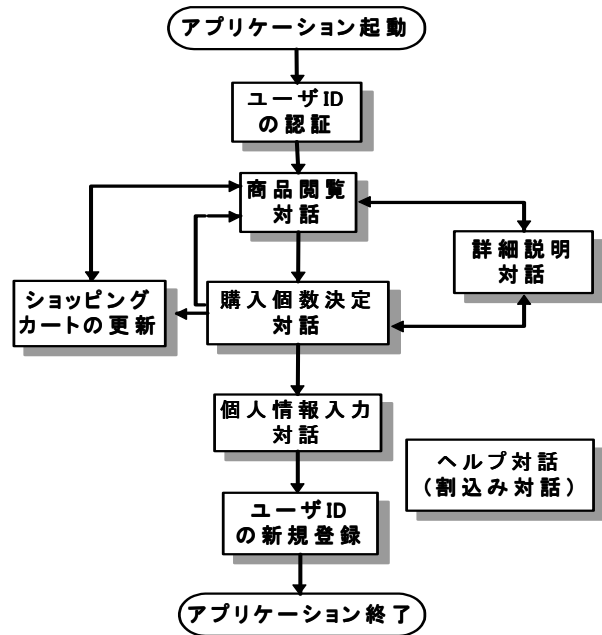


図 5 . OLS における対話の流れ

示内容と商品データに関する XML コンテンツ,表示スタイルを与える XSL スタイルシートから構成される.

図 5 にアプリケーションの大まかな流れを示す.アプリケーションを起動すると,まずユーザの認証が行なわれ,その後,商品閲覧対話に移行する.商品閲覧対話では,ショッピングカートの内容変更,購入個数決定対話への遷移や,商品の詳細説明対話への遷移が可能である.購入個数決定対話では,購入する商品の個数を決定し,個人情報入力対話か,もしくは商品閲覧対話に遷移する.また,商品の詳細説明対話,ショッピングカートの内容変更対話への移行も可能である.商品の詳細説明対話では,商品の説明終了後,遷移元のページに戻る.個人情報入力対話では,購入情報を入力した後, ID を持っていないユーザはユーザ ID 登録対話で登録をし,アプリケーションは終了する.なお,ヘルプ対話は,アプリケーション起動中の任意の時点で呼び出すことができる.

4.2. PC 向け OLS アプリケーション

PC 向け OLS アプリケーションでは,図 5



図 6 . PC 向けアプリケーションの商品閲覧ページ

に示した対話の流れのうち ,ショッピングカートの更新を除く各対話のページが画面に表示される . ユーザは各対話において音声入力 , タッチ入力 , もしくはその両方を用いて操作を行う . システムは Web ブラウザによるページの表示に加えて , 擬人化エージェントによる各種案内を行う . ショッピングカートは , 図 6- に示すように商品閲覧のページに表示され , その内容を常時変更することが可能となっている . 加えて商品閲覧ページでは , 図 6- 部分をクリック , もしくは商品カテゴリを発話することにより , 図 6- に表示されている商品一覧を変更することが可能である . なお , ヘルプページは , アプリケーション起動中の任意の時点で呼び出すことができる割込み対話として実装されている .

4.3. PDA 向け OLS アプリケーション

PDA 向け OLS アプリケーションが用いるコンテンツは , PC 向けアプリケーションと共通で , 表示画面もほとんど PC 向けアプリケーションと同様である . ただし図 7 に示すように , ショッピングカートは画面の制約上 , 商品閲覧ページに表示されず , カート更新のためのページへのリンクとして表示されるようになってい . PDA にはキーボードが備えられていないため , 文字入力は音声 , もしくはソフトキーボードを用いる . 出力は Web ブラウザ , およ



図 7 . PDA 向けアプリケーションの商品閲覧ページ

び予め wave ファイルとして保存した合成音声を用いている .

4.4. 電話向け OLS アプリケーション

電話端末では , 図 8 に示すようにユーザが商品カタログを見ながら注文を行うことを想定している . 画面出力が不可能なため , ユーザとシステムの対話は音声と DTMF を用いる . 対話の流れは PC , および PDA 向け OLS アプリケーションと同様であるが , ショッピングカートの更新は割込み対話として実装されている . また DTMF は , 商品の選択 , 個数の決定など , 数値入力の場面で , 音声とのマルチモーダル入力 (択一入力) が可能になっている . なお , 現時点で商品詳細説明とヘルプ対話が未実装である .



図 8 . 電話向けアプリケーションの利用例

5. まとめ

PC, PDA, 電話を端末とする MMI システムを実装するとともに, これらのシステム上で動作する OLS アプリケーションを開発した. 今回実装したシステムは XISL を実行する汎用システムであり, XISL およびその他文書を作成することで, 多様な Web アプリケーションを実行することができる. 我々のグループでは, OLS アプリケーションの他, 航空チケット予約アプリケーションを各 MMI システム向けに開発しており, これらも同様, 良好に動作することを確認している.

今後は新たな端末の導入, 新たなアプリケーションの作成を検討するとともに, XISL の高機能化[10], プロトタイピングツール[11]の充実に進めていきたい.

参考文献

- [1] <http://www.w3.org/2002/mmi/>
- [2] <http://www.saltforum.org/>
- [3] <http://www.w3.org/TR/xhtml+voice/>
- [4] 植田喜代志, 秋田祥史, 荒木雅弘, 西本卓也, 新美康永: "VoiceXML のマルチモーダル化の検討", 情報処理学会研究報告 2001-SLP-38, pp.43-48 (2001).
- [5] 中村有作, 小林聡, 桂田浩一, 新田恒雄: "XISL: コンテンツ記述とインタラクション記述分離の試み", 情報処理学会第 62 回全国大会講演論文集(分冊 4), pp.71-72 (2001).
- [6] K. Katsurada, H. Yamada, Y. Nakamura, S. Kobayashi and T. Nitta: "XISL: A Devices/Contents Independent MMI Description Language," Proc. of ISCA Tutorial and Research Workshop Multi-Modal Dialogue in Mobile Environments, pp.88-90 (2002).
- [7] <http://www.vox.tutkie.tut.ac.jp/XISL/>
- [8] K. Katsurada, Y. Otani, Y. Nakamura, S. Kobayashi, H. Yamada and T. Nitta: "A modality-independent MMI system architecture," Proc. of ICSLP2002, pp. 2549-

2552 (2002).

- [9] 伊勢路真吾, 福田隆, 桂田浩一, 新田恒雄: "0-gram 汎用 LVCSR と音素弁別特徴ベクトルを利用した対話音声認識の検討", 情報処理学会研究報告 2001-SLP-44, pp.213-218 (2002).
- [10] 中村有作, 桂田浩一, 山田博文, 新田恒雄: "MMI 記述言語の標準化動向と XISL の対応について", 情報処理学会研究報告 2001-SLP-44, pp.237-242 (2002).
- [11] 足立裕秋, 桂田浩一, 山田博文, 新田恒雄: "MMI システム構築のためのプロトタイピングツールの開発", 情報処理学会研究報告 2001-SLP-43, pp.7-12 (2002).