

移植性に優れた MMI システムアーキテクチャの検討

5 Q-1

大谷 佳彦[†] 桂田 浩一[†] 山田 博文[‡] 小林 聡[‡] 新田恒雄[†]

[†]豊橋技術科学大学大学院 工学研究科 [‡]豊橋技術科学大学 工学部

1. はじめに

マルチモーダル対話(MMI)は,複数のモダリティを連携させることで,自然で分かり易い対話を実現することを目的にしている。しかし,多種多様な端末から Web などのサービスを利用する場合には,端末環境に応じて対話シナリオを用意しなければならない。我々は,XML コンテンツ,対応するスタイル,および対話シナリオの三つを分離記述する枠組みを提供することにより,これら三つのドキュメントを組み合わせることで再利用することを目指している。ここで,スタイルには XSL,対話シナリオには XISL(Extensible Interaction-Sheet Language)^{[1][2]}を用いる。

本報告では,上記三つのドキュメントを利用した MMI システムアーキテクチャを提案する。提案アーキテクチャでは,移植性を高めることを目標にモジュール分割を行い,中心となる対話制御部は,端末に依存しない設計を心掛けた。以下では,XISL について説明した後,MMI システムアーキテクチャ,および提案アーキテクチャに基づき開発中の MMI システムについて述べる。

2. マルチモーダル対話記述言語 XISL

XISL は,XML コンテンツに対する MMI を記述する言語である。XISL では,複数の dialog から対話シナリオを構成する。各 dialog は,対話の最小単位を表す exchange から構成される。各 exchange は,入力記述部(operation)とアクション記述部(action)をそれぞれ一つずつ持つ。単一の入出力は,input と output により表される。また,operation や action 内では,複合した入出力を制御するタグが用意されている(sequential:逐次的,parallel:同時並列的,alternative:択一的)。図1に簡単な XISL の記述例を示す。

3. MMI システムアーキテクチャ

図2に提案する MMI システムアーキテクチャを示す。システムは,端末に依存するフロントエンド,依存しない対話制御部,および WWW 上のドキュメントサーバから構成される。

```
<?xml version="1.0" encoding="Shift-JIS"?>
<xisl>
<body>
  <dialog id="order" type="normal" comb="seq">
    <exchange>
      <operation comb="alt">
        <par_input>
          <input type="speech" event="This burger [please]"/>
          <input type="touch" match="fig[@id=ham]"/>
        </par_input>
        <input type="touch" match="button[@id=ham]"/>
      </operation>
      <action comb="seq">
        <output type="agent" event="speech">…</output>
      </action>
    </exchange>
  </dialog>
</body>
</xisl>
```

図1 XISL の例 (ファーストフード販売)

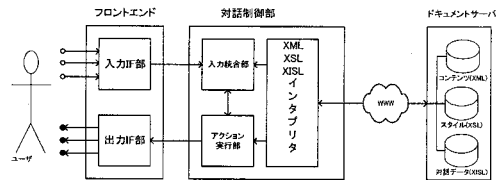


図2 MMI システムアーキテクチャ

3.1 フロントエンド

フロントエンドは,端末環境により異なる構成を持つ。入力インタフェース(入力 IF)部と出力インタフェース(出力 IF)部は,XISL の入出力記述(input, output)を理解し,モダリティに応じた入出力処理を行う。MMI システムの開発者は,端末に対応したフロントエンド構成と,XISL における入出力の記述形式を最初に定める必要がある。

3.2 対話制御部

対話制御部は,入力統合や XISL のアクション部の実行などを行う。その際,処理は実行環境や XML コンテンツに依存しないことが望ましい。以下に,対話制御部を構成するモジュールを説明する。

Highly Portable Architecture of Multimodal Interaction System: Y.Ootani[†], K.Katsurada[†], H.Yamada[‡], S.Kobayashi[‡], T.Nitta[†]; [†] Graduate School of Eng., Toyohashi University of Technology, [‡] Department of Eng., Toyohashi University of Technology; 1-1 Hibariga-oka, Tempaku, Toyohashi 441-8580, Japan

・XISLインタプリタ：XISLインタプリタは、XISLドキュメントを読み込むと共に、入力記述部とアクション記述部を切り出して各々入力統合部と出力制御部に渡す。

・入力統合部：入力統合部は、入力記述部から統合表を生成する。また、フロントエンドの入力IF部から送られる入力を受け、統合表を用いてマルチモーダル入力統合処理を行う。統合表には、入力状況を表す状態と、各状態で次の入力が生じた際の処理(状態遷移、入力受理など)が記述されている。統合表は、入力制御ブロック(dialog, exchange, seq_input など)ごとに生成される。

図3に統合表の例を示す。図中のXISLは、オブジェクトobj上でのダブルクリックの入力①、もしくはobjに対するクリックと音声"select"の逐次入力②→③が可能な対話であることを示す。統合表Aには、最初に入力可能な①と②の入力が行われたときの処理が展開されている。また、統合表Bでは、②が入力された後の処理が記述されている。状態SA-0において①が入力された場合は、その時点で入力が受理され、実行すべきアクションを持つ対話(exch1)をアクション実行部に通知する。一方、②が入力されると、統合表Bの状態SB-0へ遷移し、状態がさらにSB-1に遷移して、入力を待ち受ける。続けて③が入力されると、②→③の入力が受理され、実行するアクションが決定する。一回の入力後、一定時間入力が行われなければ、入力統合部はタイムアウトイベントを発生させる。これにより、統合処理が中止され、初期状態に遷移する。

・アクション実行部：XISLのアクション記述部には、

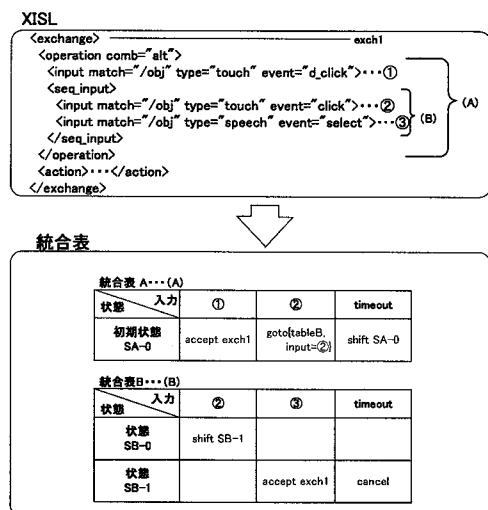


図3 入力統合表

ユーザへの出力命令(モダリティと動作パラメータなど)やXMLドキュメントへのアクセス、数値演算などが記述されている。アクション実行部は、アクション記述内容に応じて、出力IF部やドキュメントサーバと通信し処理を実行する。

3.3 ドキュメントサーバ

WWW上のサーバには、XMLコンテンツ、XSL、XISLのドキュメント群が保持され、端末環境に応じたドキュメントセットが利用される。

4. 実証システム

前節で述べたMMIアーキテクチャに基づき、実証システムをPC上に構築した。周辺デバイスとしては、タッチパネル(ディスプレイ)、マウス、キーボード、マイク、スピーカーを持つ。アプリケーションとしては衣用品のオンラインショッピングサイトを作成した。

実証システムにおいて利用する入出力モダリティ、および各モダリティで受け付け可能な入力イベントと出力動作を表1に示す。各モダリティに固有の処理は、入力IF部、出力IF部の中に作成したモジュールで実行する。

表1 実証システムで利用できるモダリティ

	モダリティ	イベント(入力)、動作(出力)
入力モダリティ	ポインティング	クリック、ダブルクリック、移動、など
	音声	認識、リジェクト
	キーボード	キーの上げ下げ
出力モダリティ	音声合成	テキストの読み上げ
	ブラウザ	MMIコンテンツ、またはWebコンテンツの表示
	動画	読み込み、再生、一時停止、停止など
	サウンド	読み込み、再生、一時停止、停止など
擬人化エージェント	アクション(アニメーション)、スピーチ(音声合成)、パルーン(テキスト表示)	

5. まとめ

本報告では、XISLドキュメントを実行するMMIシステムアーキテクチャについて説明した。本アーキテクチャの核となる対話制御部は、端末に依存せず、またモダリティの拡張にも対応できるように設計した。今後は、様々な端末への移植を通してシームレスなサービスの提供に有効であることを実証したい。

参考文献

- [1] 小林 他：“マルチモーダル対話記述言語XISLの提案”，情報処理学会研究報告 2001-SLP-37, pp.43-48 (2001)
- [2] 山田 他：4Y-06 “XISL:マルチモーダル対話記述言語の提案”，情処63回全大(2001)