

マルチモーダル対話作成支援ツール Galatea-IB の機能強化

佐藤邦俊* 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
〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1 - 1
Email: sato@vox.tutkie.tut.ac.jp

あらまし：我々は、情報処理振興協会(IPA)擬人化音声対話エージェント基本ソフトウェアプロジェクト(Galatea プロジェクト)の一部として、マルチモーダル対話システムのプロトタイピングツール Galatea-IB を開発してきた。Galatea-IB は、マルチモーダル対話システムで用いられる対話シナリオを、対話部品の組み合わせで容易に作成するための GUI ツールである。しかし、対話シナリオに演算や条件分岐、プロンプトの出力を埋め込むための対話部品がないことや、対話シナリオの動作テストを行う機能がないという問題があった。本報告では、それらに対処するために、新しい対話部品として、演算処理を行う assign、条件分岐処理を行う if、ユーザにプロンプトを出力するための prompt・reprompt を追加した。また、動作テストを行うためのテストモード機能を追加した。これにより、作成可能な対話シナリオの幅を広げ、動作テストを行うための手順を簡略化して簡単にテストを行うことができるようになった。

キーワード： マルチモーダル対話, XISL, プロトタイピングツール, Galatea プロジェクト

On Advanced Function of MMI Prototyping Tool Galatea-IB

Kunitoshi SATO*, Kouichi KATSURADA*, Hirobumi YAMADA ** and Tsuneo NITTA*

*Graduate School of Technology, Toyohashi University of Technology
** Multimedia Center, Toyohashi University of Technology

1-1 Hibarigaoka, Tempaku-cho, Toyohashi 441-8580, JAPAN
Email: sato@vox.tutkie.tut.ac.jp

Abstract : This paper presents some advanced functions of Galatea-IB, a prototyping tool for the Galatea multimodal interaction systems. We have developed it as a part of a Galatea project, which aims at developing a software toolkit for building anthropomorphic spoken dialogue agents. Galatea-IB is a GUI tool to create interaction scenarios by combining interaction parts such as input parts, output parts, and control parts. It already has basic functions to create interaction scenarios, however, it does not provide some advanced interaction parts that enable to execute arithmetic operations, conditional branches, and prompts. Moreover, it does not have a function to test interaction scenarios. With these backgrounds, we introduce some additional interaction parts into the Galatea-IB such as "assign" which executes arithmetic operations, "if" to perform conditional branches, and "prompt" and "reprompt" for outputting prompts for users. Additionally, we introduce a test-mode for testing the interaction scenarios. These advanced functions will make the work to create interaction scenarios simpler and more efficient than the work using the previous version of Galatea-IB.

Key words : multi-modal interaction, XISL, prototyping tool, Galatea Project

1. はじめに

我々は、情報処理振興協会(IPA)の擬人化音声対話エージェントプロジェクト(Galatea プロジェクト)[1]が開発した顔合成モジュール、音声合成モジュール、音声認識モジュールを用いたマルチモーダル対話システム(Galatea-MMI システム)を構築するとともに、Galatea-MMI システムで動作するアプリケーションの作成支援ツールとして Galatea Interaction Builder(Galatea-IB)を開発してきた[2]。Galatea-IB は対話部品(入力部品、出力部品、制御部品等)の組み合わせによってマルチモーダル対話シナリオを作成することを目的とした GUI ベースのプロトタイピングツールである。Galatea-IB によってモダリティの属性値(音声認識文法など)の設定や、対話シナリオの試作が、容易になったため、効率的な対話シナリオ作成が可能となった。しかし、これまでの Galatea-IB は代入処理やプロンプトの出力等、対話に頻出すると思われる要素を対話シナリオに付加する機能がなかった。また、作成中の対話シナリオの動作テスト機能がなかったため、動作テストの度に別途 Galatea-MMI システムを起動して動作させる必要があった。

今回のバージョンでは、これらの問題に対処するために、シナリオ作成用の新たな対話部品として、代入処理を行うための assign、条件分岐処理を行うための if、ユーザへの入力を求めるための prompt、および reprompt を追加した。また作成中の対話シナリオの動作確認を行うテストモードの機能を Galatea-IB 内部に追加した。これらの機能によってより効率的な対話シナリオの作成が可能になると期待できる。

以下では、始めに Galatea-MMI システムの概要、IB の機能について述べる。続いて、今回 IB に追加した新機能について述べる。

2. Galatea-MMI システム

2.1 Galatea-MMI システムの概要

図2に、Galatea-MMI システムのアーキテクチャを示す。システムは、フロントエンド、対話制御部、WWW 上のドキュメントサーバから構成される。ドキュメントサーバには、対話シナリオ、及び表示コンテンツがそれぞれ XISL[3][4]、HTML ファイルとして、保持される。これらのドキュメントのうち XISL は対話制御部に、表示コンテンツ

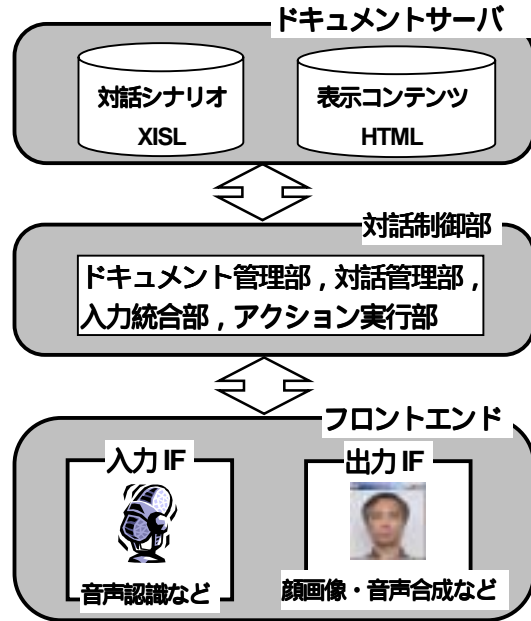


図1 Galatea-MMI システムのアーキテクチャ

表1 IPA-MMI システムで利用可能なモダリティ

	モダリティ	受け付け可能なイベント または出力動作
入力	マウス	クリック、ダブルクリック等
	音声	音声認識結果
	キーボード	キー入力
出力	ウィンドウ	コンテンツの表示
	音声合成	テキストの読み上げ
	顔画像合成	顔モデルの変更、表情、リップシンク等

はフロントエンドに渡される。対話制御部は、対話シナリオの解釈・実行、フロントエンドからの入力情報の処理、フロントエンドへの出力命令の送信を行う。フロントエンドは、Galatea プロジェクトで開発した音声入力・合成モジュール、顔合成モジュールを利用した入出力インタフェースを備え、表1に示すモダリティをサポートしている。入力インタフェースがユーザからの入力を受け付けると、その内容を対話制御部に送信し、逆に、対話制御部からの出力命令を受けると、出力インタフェースがユーザへの出力を行う。

Galatea-MMI システムは、PC 上での動作を前提に開発されている。これまでの Galatea-MMI システムは、対話制御部の動作が非常に重く、システムを動作させる場合には、フロントエンドと対話制御部を別々の PC 上で動作させる必要があ

った。しかし、その後の対話制御部の改良で、現時点では1台のPC端末上で動作が可能となっている。

Galatea プロジェクトでは、音声入力・合成、顔合成の各モジュールが分散した環境下でも動作するよう、エージェントマネージャが用意されているが、Galatea-IB では、同一システム上にすべてのモジュールを置き、Windows 上で動作することを前提に開発している。

2.2 フロントエンドモジュール

2.2.1 音声認識モジュール

音声認識エンジンには、IPA の日本語ディクテーション基本ソフトウェアに含まれる Julian を用いている。これは、京都大学音声メディア研究室で開発された有限状態文法(DFA)に基づく連続音声認識エンジンである[5]。

2.2.2 顔画像合成モジュール

Galatea プロジェクトで開発した顔画像合成モジュール[6]の Windows 版を使用している。このモジュールは、任意の顔画像と、予め用意されたワイヤーフレームモデルとの整合を行うことにより、顔モデルを生成する。表情とその継続時間の設定、および音声合成との同期をとるリップシンク等を行うことができる。また、マシンスペックに応じて FrameRate を変更することができる。

2.2.3 音声合成モジュール

合成エンジンには、Galatea プロジェクトで開発した GalateaTalk (Gtalk) [7]を使用している。Gtalk は発話のテキストを 茶筌を用いて解析し、出力された形態素情報から音声波形を生成する。

2.3 MMI 記述言語 XISL

Galatea-MMI システムはユーザとシステムとの対話の記述に MMI 記述言語 XISL(Extensible Interaction Scenario Language)を使用している。XISL は XML ベースの言語[3][4]であり、対話シナリオの可読性が高く、入出力の制御や割り込み対話の記述が容易であるという特徴を持つ。したがって XISL の利用により、システム開発者は MMI 機能を持つ Web サービスを容易に提供できる。

図 2 に XISL の記述例を示す。XISL は、一組の対話シナリオを表す dialog の集合から構成される。

```
hello.xisl
<?xml version="1.0" encoding="Shift-JIS" ?>
<!DOCTYPE xisl SYSTEM "xisl.dtd" >
<xisl version="1.0" >
  <head>...</head>
  <body>
    <dialog id="hello World" >
      <exchange>
        <operation target="hello.xml" >
          <input type="touch" event="click"
            match="/page" >
        </operation>
        <action>
          <output type="speech" event="tts-speech" >
            <![CDATA[
              <param name="speech-text" >
                Hello World!
              </param>
            ]]>
          </output>
        </action>
      </exchange>
    </dialog>
  </body>
</xisl>
```

図 2 XISL の記述例

各 dialog は、対話の最小単位である exchange の集合を内部に含み、各 exchange は、operation と action をそれぞれ一つずつ持つ。operation はユーザの入力記述部で、例えば音声やポインティングの入力待ち受けが内部の input に記述される。一方、action はシステムのアクション記述部で、合成音声や擬人化エージェントによる応答を内部の output に記述することができる。詳細については文献[4]の Web サイトを参照されたい。

3. Interaction Builder の概要

これまで音声認識、顔画像合成、音声合成モダリティの利用、および前述の XISL の各要素に対応した対話部品の付加が可能な Galatea-IB を開発してきた。以下では、IB の画面例、および従来の機能を先について説明する。

3.1 シナリオビュー

対話の流れを図式的に表示するためのシナリオビューを図 3- に示す。シナリオビューには対話の枠組みや対話の遷移、受け付ける入力イベントやそれに対する出力動作が、それぞれ一つの部品として表わされ、対話の流れに沿って配置され

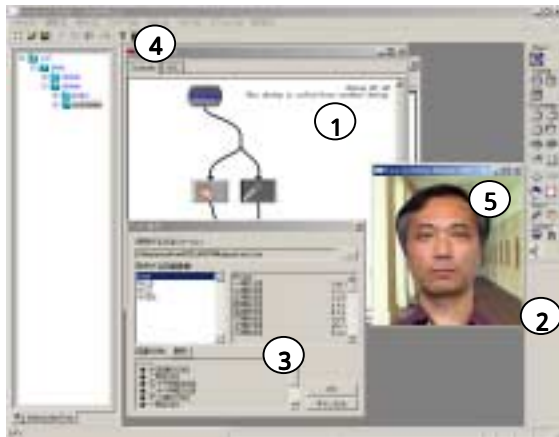


図3 IBの実行画面例

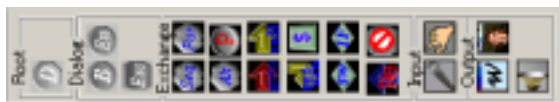


図4 対話部品バー



図5 動作確認バー

る。これにより、エディタで XISL を記述するのに比べ、構築中の対話の流れを容易に理解することが可能になった。

3.2 対話部品バー

図3- に示すバーは、対話を作成する際に利用する対話部品を並べたツールバーである。図4にその拡大図を示す。ツールバーは使用できる部品が階層ごとに分けられている。例えば dialog という枠内にある部品 begin exchange end は dialog の階層内でのみ使用可能ということを示している。root(この場合の root はツリーの最上部のことを指す)、dialog、exchange(begin・end)内にある対話部品は、対話を制御する部品である。一方、input、output の枠内にある対話部品は、入出力モダリティの属性を設定する部品である。これらの各対話部品は、同名の XISL のタグにそれぞれ対応する。ツールバーの中で、使用できない部品・ボタンは反転表示され、選択できる部品を容易に識別できるようになっている。

対話シナリオを作成する際には、ツールバーが

ら目的の対話部品を選択し、シナリオビューに drag & drop すればよい。シナリオビューに対話部品が置かれると、図3- に示すダイアログボックスが表示され、属性値など必要な情報の入力が促される。

3.3 動作確認バー

図3- に示すバーは、エージェント(図3-)の動作確認をするためのエージェントエンジンの起動や作成中の対話シナリオの動作確認をするためのテストモードに移行するボタンである。拡大図を図5に示す。エージェントの動作を確認する場合は左のボタンを、対話シナリオの動作テストを行う場合は右のボタンを押す。なお、真中のボタンは音声合成のエンジン起動ボタンでエージェントエンジンが起動した時に同時に起動する。テストモードについては4.2節で詳しく述べる。

4. 新しい機能

4.1. 対話部品の追加

IB で利用可能な対話部品に以下の4つの XISL タグに対応するものを追加した。各部品の説明は次の節以降で述べる。

- ・ <assign>
- ・ <if>
- ・ <prompt>
- ・ <reprompt>

4.1.1 assign の追加

assign は演算結果を指定した変数に代入する部品である。代入値として、文字列と数値が使用できる。XISL では文字列と数値を併用することができるが IB では、ユーザの入力間違いを防ぐために排他使用になっている。

4.1.2 if の追加

これまでの IB では、条件分岐処理に switch を使用していた。しかし、switch は対象となる値が分岐条件値と等しい時に分岐する処理で、比較演算の結果による分岐が出来なかった。そこで条件分岐処理を行う if を新たに追加することによって比較演算の処理を可能にした。if はまた、switch に比べて条件の記述を容易にする。例えば、航空チケット予約のアプリケーションにおいて、入力された便があれば条件分岐するような場合を考え

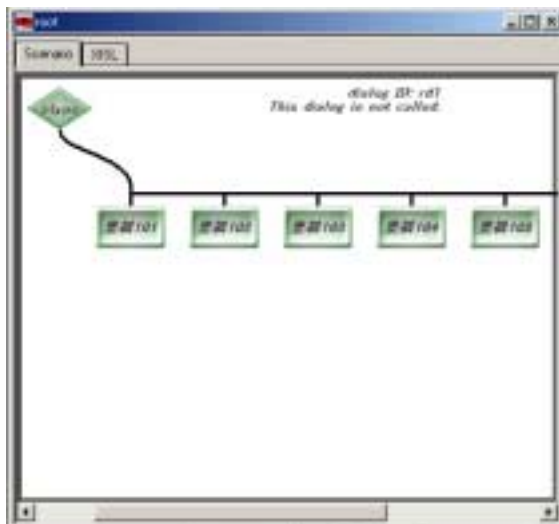


図6 switch で条件分岐

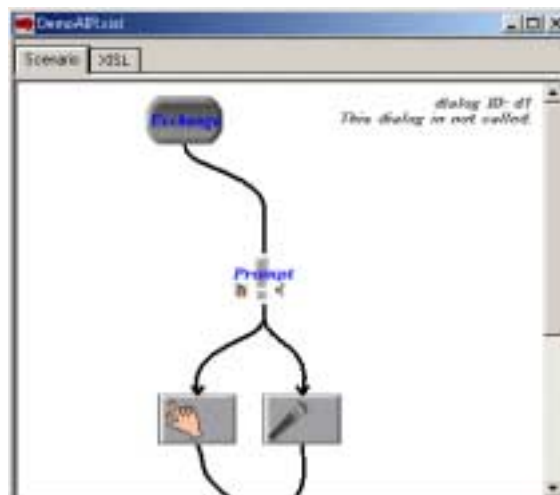


図8 prompt 部品の追加

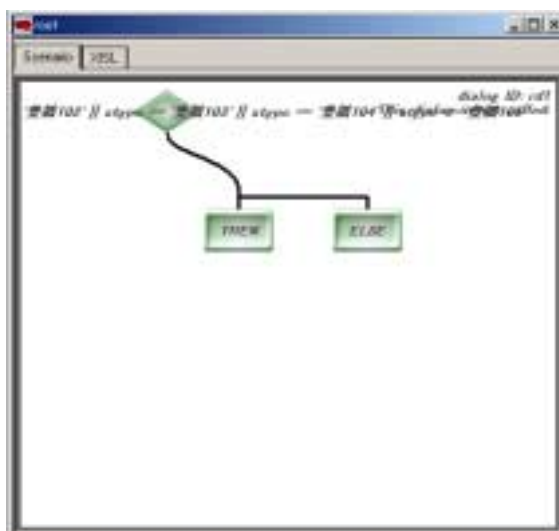


図7 if での分岐条件

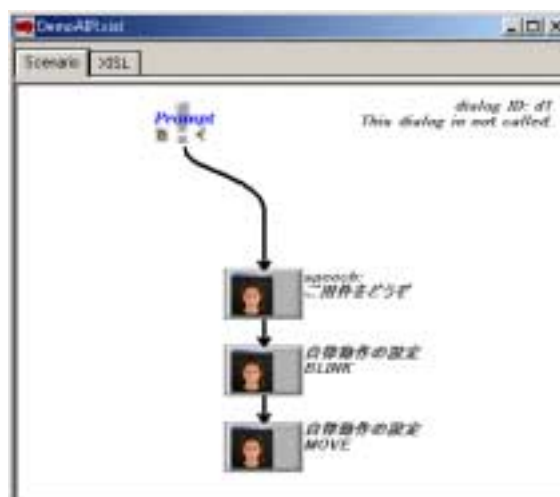


図9 prompt を利用したユーザへの入力促し

る。switch では条件値の指定に case を用いるため、図6に示すように条件の個数分 case を作成する必要があった。しかし、if では論理演算によって条件を一つにまとめて処理することができるため、図7のように分岐を“then”と“else”だけにすることができ、条件の記述が簡潔になった。なお if の条件値で扱える値は内部処理の都合上、assign とは異なり、文字列と数値の排他使用していない。

4.1.3 prompt と reprompt 追加

これまでの IB ではユーザにプロンプトを出力

する時にはbeginの階層内でoutputを用いて行う必要があった。しかし、この方法では、promptに対応するユーザ入力別のexchangeに記述されることになり、プロンプトに対応する入力という一体性が表せないという問題があった。そこで、exchangeでプロンプトを出力できるようにpromptを追加した(図8・9)。promptでは、システムの出力を途中で止めてユーザ入力を許可するバージョンと、promptの再実行を行うまでの入力待ち時間を指定するタイムアウトの設定が可能である。タイムアウトの設定とは別に、シナリオの任意の場所でpromptを再実行可能にするため

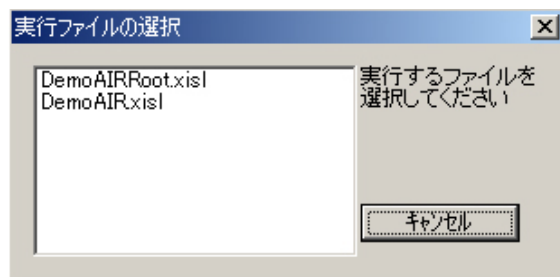


図 1 0 実行シナリオの選択

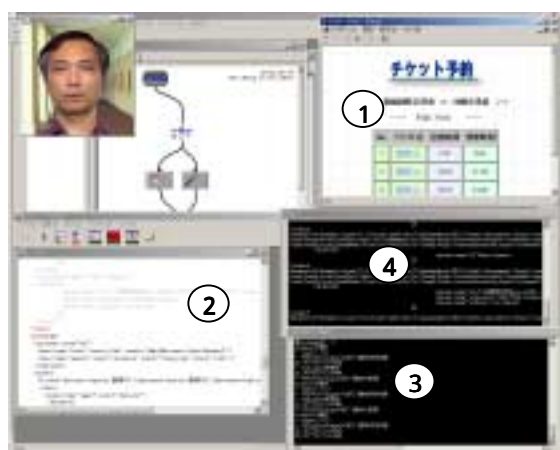


図 1 1 テストモード実行画面

に, reprompt という対話部品も導入した.

4.2 テストモードの追加

IB で作成途中のマルチモーダル対話シナリオを実際に実行して確認するために, テストモードを追加した. これまでは動作確認を行うには, 手動で Galatea-MMI システムのフロントエンドと対話制御部を起動させて, テスト対象の対話シナリオの URI を指定する必要があった. 今回追加したテストモードでは, フロントエンドと対話制御部をモジュールとして自動的に呼び出し, 対話シナリオを実行する. まず, 動作確認バーからテストモードのボタンを押すと, 図 1 0 のような実行シナリオ選択ダイアログが表示される. 実行させる対話シナリオをダブルクリックするとフロントエンド(図 1 1-) と対話制御部(図 1 1-) が起動する. その後, Galatea-IB は指定された対話シナリオの URI を自動的に対話制御部へ渡し対話を開始する. 動作の確認はフロントエンドを使用してエージェントとの対話で確認する他に, 対話制御

部の XISL から対話の流れを見て確認することもできる. また, 対話制御部にフロントエンドとの通信ログ(図 1 1-) , 対話の詳細な進行状況(図 1 1-) を表示する機能があり, もし, 動作がうまくいかない場合には, それらを使用して問題部分を見つけ出すことができる.

5 まとめ

本報告では, Galatea-IB の強化について述べた. 今回の改良では, Galatea-MMI システムで利用可能な対話部品の追加と作成中の対話シナリオをテストする機能を新たに開発した. 今回の機能強化で対話記述の幅を広げ, 動作テストの起動手順を簡略することができた. Galatea-MMI システム及び IB は情報処理進事業協会(IPA)「独創的情報技術育成事業」の支援を受けた「擬人化音声対話エージェント基本ソフトウェアの開発」プロジェクトの一部として行われており, プロジェクトの Web サイトからダウンロード可能である[1]. 今後は利用者からのバグレポートへ対応すると共に, まだ未開発になっている機能の開発を行っていきたい.

参考文献

- [1] <http://hil.t.u-tokyo.ac.jp/~galatea/galatea-jp.html>
- [2] 足立裕秋 他: “ 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 SLP-43-2, pp. 7-12(2002)
- [3] 中村有作 他: “ 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 SLP-44-40, pp.237-242 (2002).
- [4] <http://www.vox.tutkie.tut.ac.jp/XISL/XISL.html>
- [5] 河原, 李, 小林他: “ 日本語ディクテーション基本ソフトウェア(98 年度版)の性能評価”, 情報処理学会研究報告 99-SLP-26-6, pp.43-48 (2001).
- [6] 四倉達夫, 森島繁生: “ 擬人化音声対話エージェントツールキットのための表情合成技術”, 情報処理学会研究報告 2001-SLP-40-11, Feb. (2002).
- [7] 山下洋一 他: “ マルチモーダルコミュニケーションのための音声合成プラットフォーム”, 情報処理学会研究報告 SLP-40-12 ,pp.67-72(2002).