

Galatea: 音声対話擬人化エージェント開発キット

Galatea: An Anthropomorphic Spoken Dialogue Agent Toolkit

西本卓也 ¹⁾ Takuya Nishimoto	荒木雅弘 ²⁾ Masahiro Araki	伊藤克巨 ³⁾ Katsunobu Itou	宇津呂武仁 ⁴⁾ Takehito Utsuro	甲斐充彦 ⁵⁾ Atsuhiko Kai
河口信夫 ³⁾ Nobuo Kawaguchi	河原達也 ⁴⁾ Tatsuya Kawahara	桂田浩一 ⁶⁾ Kouichi Katsurada	小林隆夫 ⁷⁾ Takao Kobayashi	嵯峨山茂樹 ¹⁾ Shigeki Sagayama
下平博 ⁸⁾ Hiroshi Shimodaira	伝康晴 ⁹⁾ Yasuharu Den	徳田恵一 ¹⁰⁾ Keiichi Tokuda	中村哲 ¹¹⁾ Satoshi Nakamura	新田恒雄 ⁶⁾ Tsuneo Nitta
坂野秀樹 ¹²⁾ Hideki Banno	広瀬啓吉 ¹⁾ Keikichi Hirose	峯松信明 ¹⁾ Nobuaki Minematsu	三村正人 ¹³⁾ Masato Mimura	森島繁生 ¹⁴⁾ Shigeo Morishima
山下洋一 ¹⁵⁾ Yoichi Yamashita	山田篤 ¹³⁾ Atsushi Yamada	四倉達夫 ¹¹⁾ Tatsuo Yotsukura	李晃伸 ¹⁶⁾ Akinobu Lee	

(1) 東京大学大学院 情報理工学系研究科

(〒113-8656 東京都文京区本郷 7-3-1 E-mail: nishi@hil.t.u-tokyo.ac.jp)

(2) 京都工織大 (3) 名大 (4) 京大 (5) 静岡大 (6) 豊橋技科大 (7) 東工大 (8) 北陸先端大 (9) 千葉大

(10) 名工大 (11) ATR (12) 和歌山大 (13) ASTEM (14) 成蹊大 (15) 立命館大 (16) 奈良先端大

1 はじめに

著者らは、音声対話技術を活用した情報処理技術のいっそうの高度な利用を目指して、2003年11月より3年間の予定で音声対話技術コンソーシアム (ISTC) を発足させた [1]。ISTC の主な活動は IPA (情報処理振興事業協会) のプロジェクト (Galatea プロジェクト) で開発された「音声対話擬人化エージェント基本ソフトウェア」の発展と拡充である。この開発キットはモジュール構成の柔軟さを考慮して設計され、マルチモーダル対話の研究開発支援プラットフォームを提供している [2]。本報告では、擬人化音声対話エージェントの開発キット Galatea の概要について述べる。なお、Galatea のダウンロード方法などの情報は下記を参照されたい。

<http://hil.t.u-tokyo.ac.jp/~galatea/>

2 Galatea の構成と特長

図 1 に開発キットの全体構成を示す。各モジュールは独立のプロセスとして設計されており、対話システムは単一 PC (Note PC (Mobile Pentium III 1.2GHz, 512MB) 上で動作確認済)、あるいは分散環境 (複数 PC による並行動作) で使用することができる。以下に各モジュールを概説する。

(1) 音声認識: Julian [3] をベースに音声対話システムで要求される (a) 文法に基づく音声認識, (b) 発話中の逐次的な認識結果出力, (c) 認識処理の動的制御 (中断, 文法の切替等) の諸機能を提供している (図 2)。

(2) 音声合成: 日本語テキスト音声合成に必要な基本機能 (形態素解析 [茶筌 [4]], 読み・アクセント付与, 韻律生成, 合成波形生成) のほか, (a) 音素継続時間長を出力

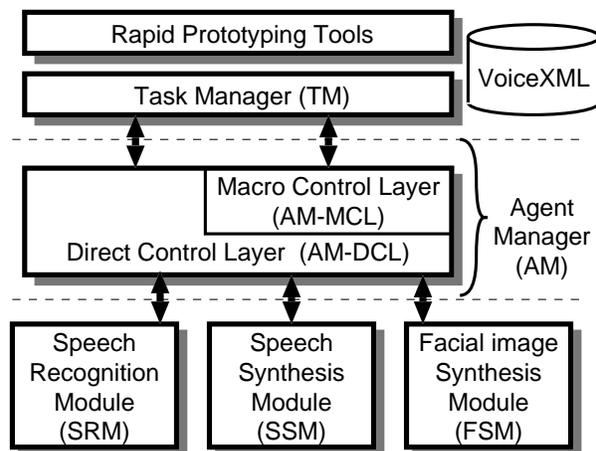


図 1: Galatea 開発キットの全体構成

し顔画像の口唇との同期が可能, (b) テキスト埋め込みタグ (JEIDA 規格準拠) による韻律制御が可能, (c) 合成音を出力途中で中断可能 (barge in 等) といった特長を持つ。合成器は HMM に基づく方式 [5] を採用し, 男女各 1 名の話者モデルを提供している (図 3)。

(3) 顔画像合成: 標準ワイヤフレームモデル中の代表点と正面写真中の対応点を, 短時間 (5-10 分) のマウス操作で整合させるだけで, 表情変化が可能である [6]。表情は怒り, 喜び, 悲しみ, 驚き, 嫌悪, 恐れ の 6 種を用意している (図 4)。音声対話のため, LipSync のほか, 自立的な動作 (うなづき, 瞬き等) を提供している。

(4) エージェントマネージャ: 対話部品が個々に規定するコマンドセットを使用して直接制御するレイヤと, 対話管理に便利なマクロコマンドを利用して制御できるレイ

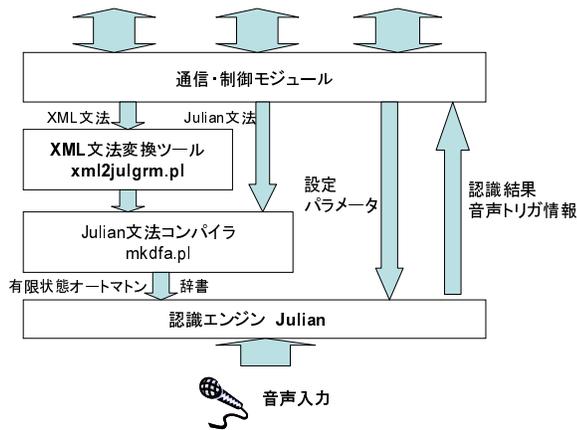


図 2: 音声認識モジュールの構成

```
<form id="main">
<field name="place">
<prompt> <emotion type="HAPPY">場所をどうぞ .
</emotion> </prompt>
<prompt count="3">
<emotion type="SAD">東京と京都のどちらですか?</emotion>
</prompt>
<grammar><rule><one-of>
<item><token sym="とうきょう">東京</token></item>
<item><token sym="きょうと">京都</token></item>
</one-of></rule></grammar>
</field>
</form>
```

図 5: VoiceXML による対話の記述例

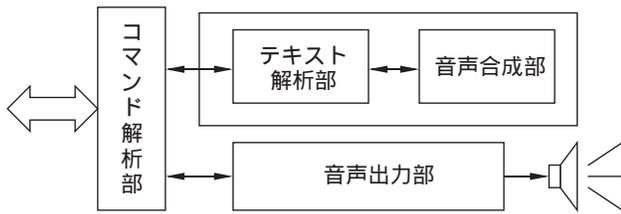


図 3: 音声合成モジュール (GalateaTalk) の構成

ヤの二つを提供している (Unix 版の場合 [6] . Windows 版では対話マネージャが各対話モジュールとソケット通信を行なう) .

3 開発支援ツール

対話記述言語として , (a) VoiceXML に GUI のためのタグを付加したもの (主に Linux 版で使用される [7] , 図 5) と , (b) モダリティの追加が可能なマルチモーダル対話向け言語 (XISL [7]; Windows 版で使用) の二つを提供している . 現在 , (a) では対話処理系および簡単な GUI ツールが , (b) ではラピッドプロトタイピングツール (Interaction Builder (IB) [8] , 図 6) が開発キットに

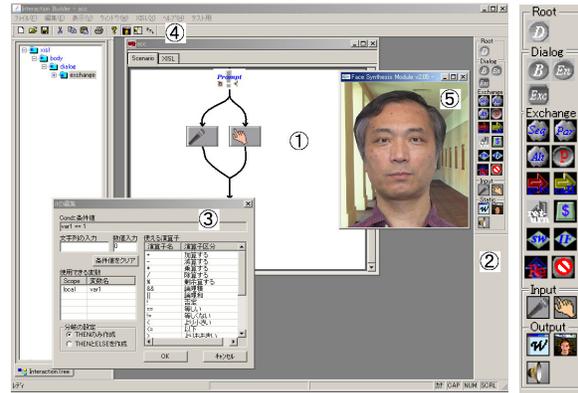


図 6: Galatea-IB の実行画面と対話部品バー

含まれている .

4 今後の予定

Galatea はオープンソース化を前提に開発された . 音声対話技術コンソーシアム (ISTC) では今後 , 各サブモジュールの改良を行なっていく予定である . また , CD-ROM 配布 , セミナー・講習会開催を通して関連研究と応用システム開発を支援する予定である .

参考文献

[1] <http://www.lang.astem.or.jp/ISTC/index.html>
 [2] 嵯峨山ほか : 情処研報 , SLP-45-10, pp.57-64 (2003).
 [3] 住吉ほか : 情処研報 , SLP-37-16, pp.91-96 (2001).
 [4] <http://chasen.aist-nara.ac.jp/>
 [5] <http://hts.ics.nitech.ac.jp/>
 [6] 川本ほか : 情処論誌 , vol.43, no.7, pp.2249-2263 (2002).
 [7] 西本ほか : 人工知能学会全大 , 2C2-04 (2003).
 [8] 足立ほか : 情処研報 , SLP-43-2, pp.7-12 (2002).



図 4: 表情合成結果の一例