

音声対話ロボット ROBISUKE による相談型対話の実現

藤江真也 坂本直宏 清水健二 宗近純一 福島健太 瀬戸祐介 小林哲則

早稲田大学 理工学部

あらまし:

相談型対話を扱う音声対話システムを音声対話ロボット ROBISUKE 上に実装した。パラ言語理解機能と、未知表現の学習機能を実装することにより、従来扱うことの難しかった相談型対話を実現するとともに、リズムある自然な対話を実現することが出来た。本稿では各機能の概要を述べる。

Realization of a Consulting Conversation on Spoken Dialogue Robot ROBISUKE

Shinya FUJIE, Naohiro SAKAMOTO, Kenji SHIMIZU, Junichi MUNECHIKA,
Kenta FUKUSHIMA, Yusuke SETO, Tetsunori KOBAYASHI

School of Science and Engineering, Waseda University

Abstract:

A spoken dialogue system for a consulting conversation is implemented on the spoken dialogue robot ROBISUKE. The para-language understanding function and the unknown expression learning function realize a consulting conversation and rhythmic and natural conversation. In this paper, we give a brief summary of these implemented functions.

1 はじめに

音声対話システムが人と自然な対話をするために必要な要素を検討し、ヒューマノイドロボット ROBISUKE 上において実装、評価を行っている。本研究では、ユーザがある事柄に関してシステムに相談をするという相談型対話を扱う対話システムを実装した。

相談型対話ではシステムの提案に対してユーザが意見を返す中で自分の要求をはっきりとさせていくという特徴を持つ。このような状況では、ユーザがシステムの提案をどう評価したかを正確に理解することが重要である。また、要求がはっきりしない状況での要求表現は、同じ内容であっても様々な形で表現される。システムにとって新規の表現で要求が行われることも多く、このような表現については適応的に学習する必要がある。

本稿では、対話ロボット ROBISUKE について紹介するとともに、相談型対話システムを実現するにあたって提案した、パラ言語を用いた発話態度の認識、ネットワークモデルを用いた要求表現の学習について述べる。

2 音声対話ロボット ROBISUKE

ROBISUKE は小型のヒューマノイドロボットで、音声対話システムのプラットフォームとして必要な

連絡先：〒169 東京都新宿区大久保 3-4-1 早稲田大学理工学部 コンピュータネットワーク工学科 小林 哲則
TEL: (03)5286-3379 FAX: (03)-3205-9381

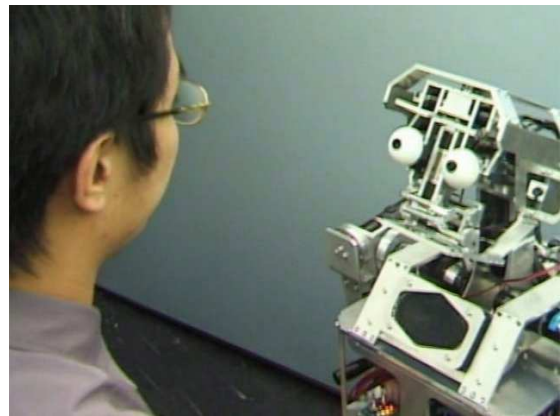


図 1: 音声対話ロボット ROBISUKE

インターフェースを備えている。本研究の目的の一つとして、対話でやり取りされる様々な情報の役割を明らかにすることが挙げられる。そのため ROBISUKE には、人間同士のコミュニケーションで重要な役割を持つ、両眼、口、眉毛、両腕などが実装されている。

実際の対話システムは、音声認識、画像認識、対話制御、モータ制御など様々なモジュールが互いに通信をすることによって対話を進行する。これらの

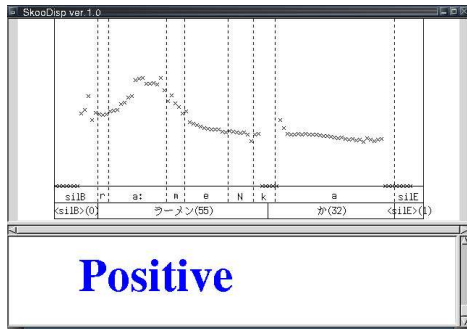


図 2: 韻律に基づく発話態度の認識の様子



図 3: 頭部ジェスチャ認識の様子

モジュールは全て、黑板モデル、publish/subscribeモデルに基づいて実装した通信アーキテクチャ上で動作する [2]。

対話中の ROBISUKE の様子を図 1 に示す。

3 パラ言語の理解

パラ言語とは、発話に含まれる言語情報以外の情報のうち、言語情報を補助する役割を担うものである。例えば、「ハンバーガーなんかどう」と提案された場合、「いい」「やだ」などとはっきり言語情報のみで提案に対する評価を理解できる発話をすることは稀で、「ハンバーガーか」と復唱するなどして、明示的な評価を提示しない場合が多い。このような発話からは、言語情報のみからでは提案に対する評価を得ることができないが、声の調子や表情などのパラ言語を併せて受け取ることで、何らかの評価を得られることも多い。パラ言語情報には様々なものがあるが、本研究では特に韻律と頭部ジェスチャに注目し、システムが提案した内容に対するユーザの評価を理解する [3]。

韻律からは、基本周波数 (F0) や話速から得られる特徴量を用いて、肯定的/否定的な態度を認識する [4][5]。図 2 に態度認識モジュールの動作画面を示す。上部が F0 のパターンと音素アライメント、下部に認識結果が表示されている。

頭部ジェスチャは、肯定的なものとしてうなづき、否定的なものとしてかしげと首振りを認識する。頭部領域のオブティカルフローを特徴量とし、HMMを用いてスポッティング認識を行う [6]。図 3 に頭部ジェスチャ認識モジュールの動作画面を示す。中央が入力画像、左部にオブティカルフロー、右部に認識結果が表示されている。

これら別々に得られた結果を統合し、提案に対する評価として解釈し、対話を制御する。

4 未知表現の学習

昼食に何を食べるか、どの店に行くかを対話で決めるためには、ユーザの要求を理解し、その要求をで

ユーザの未知表現: 「風邪を引いているんだよ」

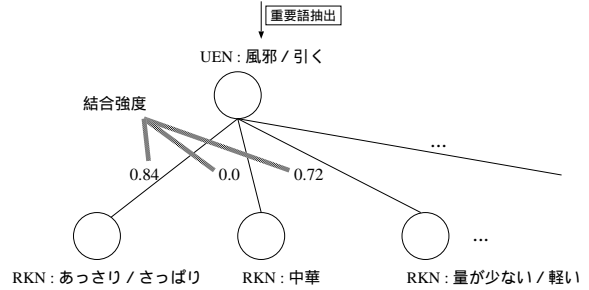


図 4: 表現と検索条件のネットワーク

きるだけ満たす提案を行うことが望ましい。ユーザの要求表現は多岐にわたり、それらを全て事前に用意しておくことは困難である。一方、検索対象 (ここでは店) の情報はそれに比べると少数であり、事前に用意することが比較的容易である。そこで、これらの関係を重み付きのネットワークで表現し、対話を行うに従って重みを変化させることで、新出の要求表現と検索対象の属性情報との関係を学習する [7]。

ネットワークの一部を図 4 に示す。UEN, RKN はそれぞれ、ユーザの要求表現に対応するノード、店の属性情報に対応するノードを表す。この例では、ユーザが「風邪を引いている」ということを伝えることにより「あっさりしたもの」「量が少ない」という属性を持つ店を要求していると理解する。

5 むすび

音声対話ロボット ROBISUKE と、その上で動作する音声対話システムについて概要を述べた。今後、このシステムをベースに、ロボット側のパラ言語の生成に関する検討や、対話システムの評価等を行っていく予定である。

参考文献

- [1] 小林哲則, “ROBISUKE:新世代の対話ロボット,” 人工知能学科学会資料, SIG-Challenge-0318-1, pp.1-6, 2003.
- [2] 松坂要佐, 於久健太郎, 小林哲則, “多機能ロボット開発のための情報共有アーキテクチャの設計と実装,” 電子情報通信学会論文誌 D-I, Vol.J86-D-I, No.5, pp.318-329, 2003.
- [3] 藤江真也, 江尻康, 菊池英明, 小林哲則, “パラ言語の理解能力を有する対話ロボット,” 情報処理学会研究技術報告, SLP-48, pp.13-20, 2003.
- [4] 八木大三, 藤江真也, 菊池英明, 小林哲則, “韻律情報を用いた肯定的/否定的態度の認識,” 日本音響学会春季研究発表会, pp.141-142, 2004.
- [5] Fujie, S., Yagi, D., Kikuchi, H., and Kobayashi, T., “Prosody based Attitude Recognition with Feature Selection and Its Application to Spoken Dialog System as Para-Linguistic Information,” Proc. of ICSLP2004, 2004.
- [6] 江尻康, 中島慶, 藤江真也, 小林哲則, “対話ロボットの動作中における頭部ジェスチャ認識,” 電子情報通信学会研究技術報告, Vol.103, No.453, PRMU2003-153, pp.13-18, 2003.
- [7] 細川健一郎, 藤江真也, 小林哲則, “検索・提案型対話システムのためのユーザとのインタラクションによる適応的意図理解,” 人工知能学会研究資料, SIG-SLUD-A302-5, pp.21-28, 2003.