人間の発話に協調的に振舞う音声対話システム（人と関わる）

研究担当者：北岡　教英（豊橋技術科学大学）
研究代表者：北岡　教英（豊橋技術科学大学）[公募研究]
研究期間：平成１５年度～平成１６年度

研 究 成 果 概 要

【目的】

ヒューマノイドロボットなどのような機械では，人間同士の会話と同様に相槌や割り込みなどの応答を返し，より円滑な対話を行うことが期待される．人間同士の対話では相槌や話者交替をタイミングよく繰り返すことによって継続されていく．本研究では，まず自然な雑談対話をする上で最も重要である相槌･話者交代のタイミングを生成する仕組みを構築する．そして，このタイミング生成器を用いた対話システムのアーキテクチャを提案し，実際に雑談対話システムを構築する．

【手法】

これらの発話タイミングは，対話相手の最後の発話の韻律的情報や表層的な言語情報をトリガーとして決定されているという知見がある．そこで，これらに基づく情報を素性とした決定木によるリアルタイムタイミング生成器を構築する．決定木は事前に学習データから C4.5 アルゴリズムで学習しておく．動作時には，ユーザのポーズ中でセグメント（100ms）ごとにシステムがとるべき動作（相槌･話者交代･自発話待ち･相手継続発話待ち）を決定する（図1）．用いた素性は，直前のユーザの発話における発話長，最後の自立語の品詞，発話区音声末100msのピッチおよびパワーの変動パターンなどの韻律的･表層的言語情報である．

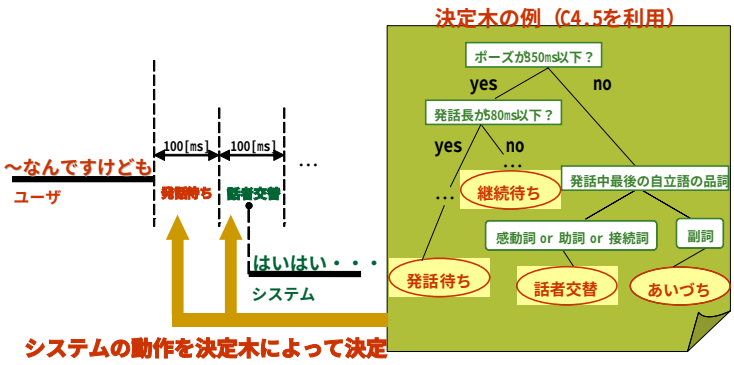

図1．決定木による応答タイミング生成

このタイミング生成を導入した対話システムのアーキテクチャを図２　に示す．入力音声は音声入力部によって特徴パラメータ抽出，ピッチとパワーの抽出が行われ，特徴量は音声認識部(300 語程度の文法駆動型，認識途中経過を出力可能)，ピッチとパワーは応答タイミング生成部へと送られる．応答タイミング生成部であいづち，話者交替，発話待ちの判定を行い，その結果に従い，応答生成部がELIZA 式に応答文を作成し実時間で応答を行う．

【結論】

実際の人間同士の対話の一方を入力とした評価により，人間のタイミングの模倣が適切にできることを確認した．また，人間とシステムとの対話を聞いて評価する評価実験により，相槌については人間同士と変わらない評価を得た．話者交代は応答内容が不十分で評価は下がるが，タイミングのみに関しては良い評価であった．さらにシステムと実際に対話して評価する実験では，対人間には及ばないものの，親しみやすく対話が自然に感じるとの感想を得た．

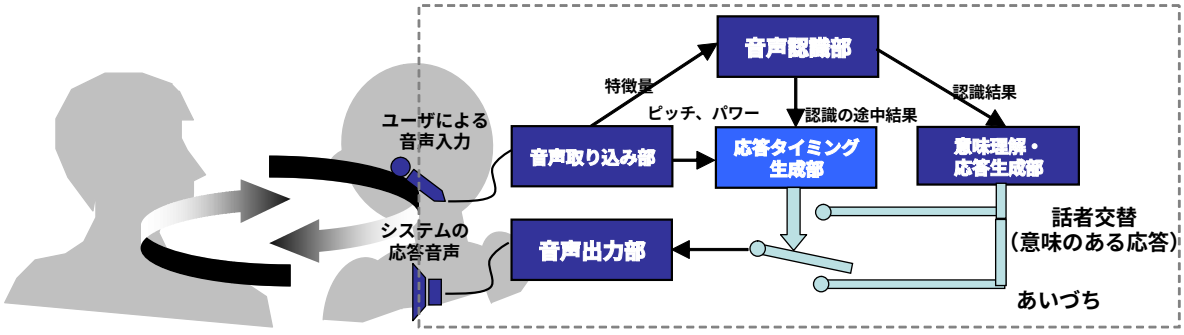

図2．応答タイミング生成部を用いた対話システムアーキテクチャ

論　文　発　表　等

1. Kitaoka et al.: Response timing detection using prosodic and linguistic information for human-friendly spoken dialog systems, Journal of The Japanese Society for Artificial Intelligence, Vol.20, No.3 SP-E, pp. 220-228, 2005.
2. Kitaoka et al.: Response timing detection using prosodic and linguistic information for human-friendly spoken dialog systems, Second International Workshop on Man-Machine Symbiotic System, pp.81-91, 2004.
3. Takeuchi et al.: Timing detection for realtime dialog systems using prosodic and linguistic information, International Conference: Speech Prosody 2004, pp. 529-532, 2004.
4. Takeuchi et al.: Generation of natural response timing using decision tree based on prosodic and linguistic information, EUROSPEECH2003, pp. 609-612, 2003.