

## 会話型ロボットを用いた介入療法における認知症進行度の評価システム

代表研究者 三重大学大学院工学研究科 助教 川中普晴  
共同研究者 鈴鹿医療科学大学 教授 山本皓二

### 【まとめ】

本研究課題では、会話型ロボットを用いた介入療法において、認知症の進行度を評価するためのシステムの構築に関する研究開発を行った。ここでは、会話型ロボットとして ifbot を使用し、日常会話から認知症の進行度を評価するためのプロトタイプシステムを構築した。また本課題では、より自然な会話のための会話誘導・返答方法について研究するとともに、高齢者音声の認識のための音響モデル構築法についても検討した。

### 1. 研究の目的

#### 1.1 社会的背景

高齢者人口の増加傾向にともない、認知症を患う高齢者の人口も増加する予測されている[1]。この傾向は社会的・医学的に問題となっており、認知症の予防・改善は必要不可欠な課題である。これらの課題を解決する方法として現在、ロボット介入療法（以下、ロボットセラピー）に関する研究が盛んに進められている。特に、昨今では簡単な会話機能を有するロボット（図 1）も開発されており、ロボットとの会話を利用した認知症の予防・改善を試みる取り組みは多い[2,3]。

一方、多くの施設では利用者の認知症進行度を評価するためのチェックテストが実施されている。このテストにより、利用者の認知症の程度（進行度）を定量的に評価できるとともに、被験者の認知症の進行度を経時変化として観察することが可能となる。そのため、定期的なテストは福祉医療において非常に重要となる。しかしながら、チェックテストの実施は、被験者がテストを意識することによる評価のゆらぎ、継続的にテストのための実施者の負担といった問題もある。そのため現在、リクリエーションやロボットの会話から認知症の進行度を評価するシステムの研究開発が求められている。

#### 1.2 研究の目的

そこで本研究では、会話型ロボットとの日常会話を利用した認知症評価システムに関する研究開発を目的とする。本課題では、これまでの研究成果を活用し、会話型ロボットを用いてプロトタイプを開発する。また本課題では、人間とロボットとのより自然な会話を実現するための返答方法の自動生成法について検討する。またさらに、各話者の話し方の特徴を考慮した音声認識のための音響モデルの構築法（追加学習法）についても検討する。

### 2. 研究方法と経過

#### 2.1 日常会話を用いた認知症評価

これまでに我々は、「会話の中に取り入れやすいこと」をコンセプトにした認知症の進行度評価法を提案するとともに、福祉施設での評価実験を行ってきた[4]。本システムでは、従来研究の成果を活用し、認知症の評価を行うこととした。

表 1 に、試作システムにて用いる高齢者への設問（設問データセット）の例を示す。認知症の評価には、時間的・地理的な状況把握能力（見当識）に加えて、聴覚性・言語性に関わる即時記憶や近似記憶も重要となる。そのため本システムでは、これらを含めた設問データセットを複数作成した。なお、これら設問データセットは、従来研究の成果に基づいて作成した。



図 1 会話型ロボット ifbot

表 1 設問データセットの一例

|                                            |   |   |
|--------------------------------------------|---|---|
| ①時間的見当識                                    |   |   |
| 「今日は何曜日ですか？」(1点)                           |   |   |
| 回答                                         |   |   |
| 「もう還暦は迎えられましたか？」(1点)                       |   |   |
| Yes□ No□                                   |   |   |
| 回答                                         |   |   |
| ②地理的見当識                                    |   |   |
| 「ここは伊勢市ですか？」(2点)                           |   |   |
| Yes□ No□                                   |   |   |
| 回答                                         |   |   |
| ③聴覚性言語性即時記憶,近時記憶                           |   |   |
| 「帰りに白菜, にんじん, ねぎを買って帰るのですが, 覚えておいてくれませんか？」 |   |   |
| 【即時】(各1点) 【近時】(各2点)                        |   |   |
| 白菜                                         | □ | □ |
| にんじん                                       | □ | □ |
| ねぎ                                         | □ | □ |

## 2.2 システムの構成

本研究では, 上述のような設問データセットを会話ロボットへ実装するため, 会話型ロボット“ifbot”を用いた. 図2に作成したシステムの概要を示す. 本システムではまず, ユーザ(高齢者)が発話した音声情報が音声認識エンジンJulius[5]に入力される. Juliusは音声データをテキストデータに変換し, その結果は設問に対する正誤判定や話題内容の認識に用いられる. テキスト化された発話内容はまず, 井筒らの方法を用いてその内容が(話題)が推定される. 推定結果を用いて, 設問データセットの質問を問いかけることができるように, 話題を誘導する応答文が生成される. なお, この時点で会話の誘導ができないと判断した場合は, 相槌や聞き直しといった汎用的応答を生成してユーザに回答することにより, 会話が中断しないようにした.

## 2.3 応答文の自動生成方法と評価実験

本システムでは, 高齢者からの発話内容を理解し, 現在の会話内容から評価のための設問

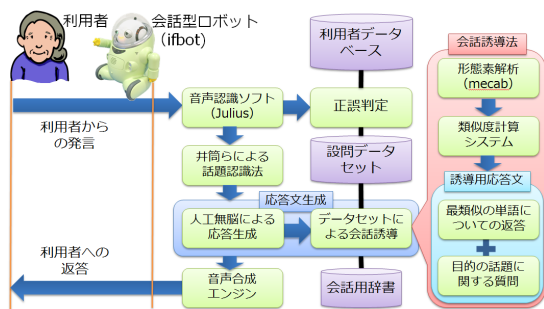


図2 システム概要図

データセットまで自然に会話を誘導することが重要となる。そこで本研究では, 応答文の生成法について検討した。

図3に会話誘導ならびに返答の自動生成方法の概要図を示す. ここではまず, 井筒らの方法[6]により, 現在の会話内容(話題)を特徴付ける概念を抽出する. 一方, 認知症評価に用いる全ての設問データセットはシステム内のデータベースに登録されており, その設問に誘導するために必要と思われる会話の話題(概念)も併せて登録されている. 本システムでは, 抽出された概念と会話誘導用に登録されている概念との類似度が最も高く, かつ予め決められたしきい値より高い場合, 会話誘導が誘発される. 誘導が誘発された場合, システムは現在の話題の下位概念と誘導先の全ての下位概念との類似度を計算する. 例えば図の例では, 入力(高齢者の発話)に現れる「雨」と「暑い」について, 誘導先の話題「季節」の下位概念である「雨季・梅雨」「夏」「収穫期」との類似度が計算される. そして, これらの中から最も類似度の高い概念の組(図では「雨」と「雨季・梅雨」の組)を選択する. 選択された単語を用いることにより, 次の返答が生成され, ユーザに投げかけられる.

評価実験では, どの程度「自然な応答文が生成されるか」検討するため, 以下のような評価実験を行った. ここでは以下の2点に注目して評価することとした.

- ① どの程度日常会話から設問データセットの質問へ会話を誘導することができるか
- ② どの程度自然な会話ができるか

実験では, 音声認識の誤認識による会話の不成立を防ぐため, 音声認識と合成音声の入出力部分を除いた会話本体部分のみを用い, テキスト形式による評価とした. また, 話題が限定された状況を実現するため, 被験者には「天気」の

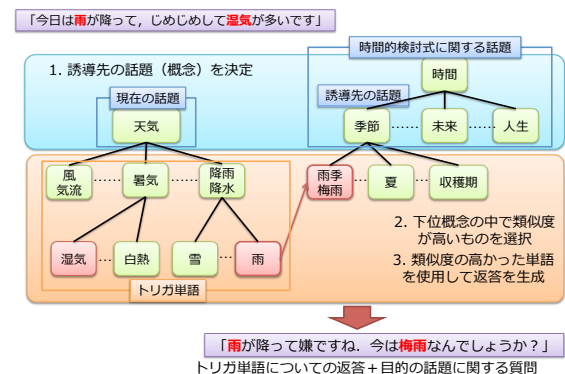


図3 返答文の自動生成方法

話題に限定してもらい、誘導先的话题を「季節」とした。5名の被験者に(a)ロボット(試作したシステム)との会話、(b)実際の人間との会話を各1回ずつ(計2回)行なってもらい、その後、5件法を用いたアンケートにより会話の自然さについて評価した。

## 2.4 音響モデルの追加学習法に関する検討

会話型ロボットが高齢者と日常会話を交わすには、マイクから得られる音声を音声認識エンジンによりテキスト化しなければならない。しかしながら、一般的な音声認識エンジンは、多数の話者音声から作成された汎用の音響モデルを用いているため、発話者固有の音響的特徴が顕著に現れる自由発話(日常会話)への適用は難しい。本課題では、この問題を解決するため、日常会話から各個人の発話の特徴と学習し、音響モデルを再構築する方法についても検討した。

実験では、国立情報学研究所音声資源コンソーシアムにて公開されている音素バランス文・語「ATR 音素バランス文 503 文」を用いた。本コーパスは、2 音素連鎖 402 種と 3 音素連鎖 223 種計 625 種をバランス良く含んでいる文のセットである。実験ではまず、音声認識エンジンを用いて音素単位での認識を行う。得られた認識結果とコーパス中のラベルと比較することにより混同行列を作成し、利用者の音素誤り傾向データを収集することとした。本稿では、音素誤り傾向の調査のため、学生を対象に以下のような実験を行った。

- ① コーパスを基に発話原稿およびラベルを作成する。
- ② 発話原稿を基に音声を録音し、音声認識エンジンを用いて音素を抽出する。
- ③ 得られた認識結果と①で作成したラベルとを比較し、音素誤り情報を取得し、混同行列を作成する。

## 3. 研究の成果

### 3.1 構築したシステム

図4に、本課題で開発した試作システムの写真を示す。ifbotの横に設置されている小型サーバには、会話内容の理解や会話誘導のモジュール、データベース等が搭載されている。ifbotの胴体内部にはLinuxベースの小型組み込みシステムが入っているがその性能は極めては低いため、ロボット表情創出以外の処理を内部コンピュータに課すことは難しい。また、設問

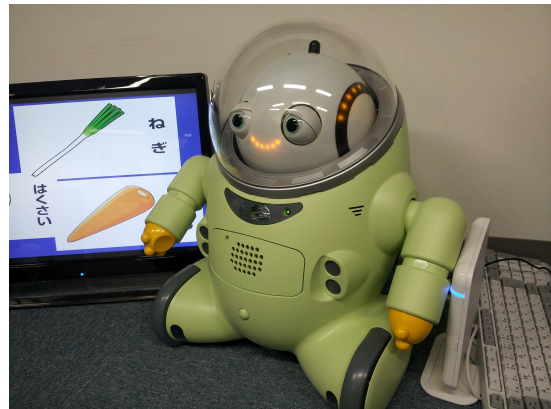


図4 開発しているシステムの概観

データセットの中にはイラスト等を表示させて回答を求める設問も取り入れていく予定であるため、横に設置されているタッチパネルディスプレイからの情報も処理しなければならない(写真ではデモ用イラストを表示)。そのため、試作システムではロボット本体とは別のサーバにこれらの情報を処理させ、サーバ側からifbotを制御する(喋らせる)構成とした。なお、現段階で一部機能に不具合が見つかっており、その修正・改良作業を進めている。試作システムの実用化の目処がたち次第、ロボット内部の組み込みシステムの高機能化と提案方式の実装を進める計画である。

### 3.2 返答文生成方法に関する検討

実験の結果、会話の誘導では(a)ロボットとの会話では5回中3回、(b)実際の人間との会話では全ての場合において質問まで誘導することができた。人間との会話においては、高い確度で会話の誘導に成功しており、かつ質問を問いかけることができると考えられる。すなわち「天気」の話題から「季節」の話題に意図的に誘導することにより、日時を題材にした評価用設問を問いかけることが可能である。また、ロボットとの会話においてもある程度会話の誘導ができていられると思われる。しかしながら、ロボットとの会話の場合「天気が悪いのはいやですね。明日も雨がふるのでしょうか?」という質問に対し、「降ると思います。」といった返答が入力された場合には、入力文中に「天気」や「季節」の話題に関する単語が存在しないため、誘導に失敗するケースが見られた。このように、話題が外れた時、あるいは非常に短い返答の場合には適切に話題を誘導することが困難である。今後は、このようなケースにおいても適切に会話誘導が行われるように、辞書の追加や文



脈の理解といったシステムも必要になると考えられる。

表 2 に、実施したアンケートの結果を示す。アンケートの結果、全ての被験者が会話の相手（ロボットか人間か）を正しく判断できていた。また、話題誘導が成功した被験者については、ほぼすべての項目で、高いスコアを示した。一方、会話の誘導に失敗した被験者については、会話に違和感を覚えたという意見が多く見られた。アンケートの結果から、今回の実験では試作システムにおけるロボットとの会話においても上手く会話を誘導するとともに、設問データセットを用いることによって「テストである」ことを意識させずに認知症の進行度を評価できると思われる。しかしながら、現段階では実験時に対象とする話題を「天気」に限定している。実環境下での運用を考えた場合、複数の話題にわたって会話が行われるため、今後は様々な話題状況下での評価実験を進めるとともに、話題の変化にも柔軟に対応できるようにシステムに拡張していく予定である。

### 3.3 音響モデルの追加学習法について

表 3 に、実験結果から作成された混同行列の一部を示す。実験の結果、個人の音素誤りの傾向に特徴があることが明らかとなった。すなわち、混同行列を用いることにより、音素誤り傾向の解候補を探索することができる可能性が示された。一方、音素認識の結果については、1 対 1 対応ではなく、他の音素に吸収されてしまうことによる音素の消失や、1 音が複数の音素に分裂して認識されるといった現象も見られた。このような認識結果に対しても対応可能な混同行列を作成することが必要であるとともに、これらの特徴を積極的に利用したモデル構築法について研究開発を進めていく必要がある。

表 2 試作システム似関するアンケート結果

| 内容             | 人間  | ロボット |
|----------------|-----|------|
| 会話内容は適切だったか    | 4.4 | 3.2  |
| 会話の流れは自然だったか   | 4.4 | 3.0  |
| 会話に集中できたか      | 4.6 | 3.6  |
| 会話が楽しかったか      | 3.8 | 3.0  |
| ロボットは親しみやすかったか | 4.2 | 3.2  |

表 3 作成された混同行列（一部）

| 正解音素 \ 認識音素 | a   | i   | u   | e  |
|-------------|-----|-----|-----|----|
| a           | 220 | 83  | 3   | 1  |
| i           | 74  | 151 | 51  | 99 |
| u           | 30  | 13  | 200 | 12 |
| e           | 26  | 72  | 5   | 35 |
| o           | 17  | 1   | 0   | 2  |

### 4. 今後の課題

今後は、現場での運用のために、設問データセットを含めたコンテンツの拡充を進めるとともに、応答文自動生成システムの更なる改良を進めていく予定である。また、本研究により得られた実験結果を活用し、高齢者の発話特徴量を積極的に用いた新たな音響モデルの追加学習方法についても研究を進めていく予定である。

### 5. 研究成果の公表方法

本研究の成果は、計測自動制御学会の支部大会、日本知能情報ファジィ学会の第 28 回ファジィシステムシンポジウム、平成 23 年度電気関係学会東海支部連合大会にて発表した。今後の予定として、計測自動制御学会支部大会にて 2 件の発表を予定しているとともに、医療情報学会や人間工学会での発表を予定している。

### 引用文献

- [1] 人口統計資料集 2012 年版, 国立社会保障・人口問題研究所, <http://www.ipss.go.jp/>
- [2] 荒木敦志, 加納政芳, 種田行男 他: 会話ロボットを用いた高齢者のための Robot Assisted Activity の試み, FIT2007, 第 3 分冊, pp.581-583, (2007)
- [3] 株式会社 ifoo (イフー), <http://www.ifoo.co.jp>
- [4] 井筒悠太, 住吉涼, 川中普晴 他, : ロボットとの会話を利用した認知症評価システムの実現に向けた検討, 医療情報学, Vol. 32, No. 2, pp.73-81, 2012
- [5] 李 晃伸, 河原 達也: Julius を用いた音声認識インタフェースの作成, ヒューマンインタフェース学会誌, Vol. 11, No.1, pp.31-38, 2009
- [6] Y. Izutsu, et. al., “A Study on Conversational Content Recognition Method Using Japanese Word Net for Robot Assisted Therapies,” J. of Advanced Computational Intelligence and Intelligent Informatics, Vol. 16, No. 1, pp. 62-68, 2012