

# 子どもを対象とする感情を持った雑談対話システム

山本 賢人<sup>†</sup> 大村 俊哉<sup>‡</sup> 根本 凌<sup>†</sup> 和田 史織<sup>‡</sup> 杉本 徹<sup>†</sup>  
 芝浦工業大学 工学部<sup>†</sup> 芝浦工業大学大学院 理工学研究科<sup>‡</sup>

## 1. はじめに

近年、雑談対話システムの研究やビジネス利用が盛んである。たとえば、NTT ドコモ社の「しゃべってコンシェル」、ソフトバンク社の「Pepper」が例として挙げられる[1]。また、子どもを対象とした対話システムの研究として「たけまるくん」が存在する[2]。

一方で子どもが対話システムを利用する場合、子どもが退屈せず楽しめるための話題や機能の充実といった課題が考えられるが、このような課題に焦点を当てた研究は少ない。

そこで、本研究ではシステムの対象年齢を 7～12 歳程度と仮定し、子どものユーザとクイズや石取りゲーム、雑談を楽しむことができるシステムの開発を目指す。クイズや石取りゲームといった子ども向けの対話の内容に加えて、システムに感情を持たせることや強化学習を行い子どもに適応することで、子どもが退屈せず利用できるようにする。

## 2. システムの概要

提案するシステムの処理の流れを図 1 に示す。

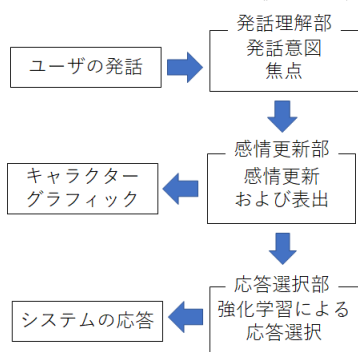


図 1: システムの処理の流れ

### 2.1. 発話意図および焦点情報の抽出

発話理解では、音声認識で認識された発話文から様々な情報を抽出する。まず、ユーザの発話から発話意図として 14 種類の発話行為タイプおよび属性を抽出する。発話行為タイプは LIBLINEAR [3] を用いてロジスティック回帰で求める。属性は名詞、形容詞、動詞等を発話行為タイプに基づき形態素解析の結果から自動的に抽出

A dialogue system with emotions for children  
<sup>†</sup>Kento Yamamoto, <sup>‡</sup>Syunya Omura, <sup>†</sup>Ryo Nemoto,  
<sup>‡</sup>Shiori Wada, <sup>†</sup>Toru Sugimoto  
<sup>†</sup>College of Engineering, Shibaura Institute of Technology  
<sup>‡</sup>Graduate School of Engineering and Science, Shibaura  
 Institute of Technology

する。一方、焦点は発話文の中心語であり、具体的には名詞およびそれがかかる述語の組み合わせと定義する。焦点は CRFsuite [4] を用いて形態素単位で焦点か否かの系列ラベリングを行い求める。

### 2.2. 感情の更新および表出

発話理解部で得られた結果を基にシステムの感情の更新を行う。本研究では、感情の概念をシステムの情緒と気分の二つに分けて定義する。情緒は瞬間的に生じる短くて強い感情である。気分は生じた情緒や内的状態に起因して生じる、弱いながらも長く持続する感情である。情緒と気分は Russel の感情円環モデル [5] に従い更新される。このモデルでは、感情は「快-不快」、「覚醒-眠気」の二次元平面上に存在し、極座標ベクトルの方向および大きさをそれぞれ感情の種類と大きさとして表現する。本研究で扱う情緒の種類を表 1 に示す。

表 1: 生起する情緒の種類

	覚醒	眠気
快	喜び	平常
不快	怒り	悲しみ

生起する情緒は「快-不快」、「覚醒-眠気」の正負により決定する。「快-不快」は、ユーザ発話文に含まれる名詞、動詞、形容詞、副詞の好感度の平均を用いる。好感度データには、単語感情極性対応表 [6] を用いる。この表は、ある単語が一般的に良い印象を持つか、悪い印象を持つかを -1 から +1 の実数値で表したものである。「覚醒-眠気」はシステムがユーザの話に興味を持つかどうかで判定する。「ユーザの発話行為タイプが陳述ならばスコアを上げる」などのルールに従いスコア付けを行い、値を決定する。

情緒と気分を以下の方法で更新する。まず上記の手法で「快-不快」の好感度  $Fr$  および「覚醒-眠気」の覚醒度  $Ar$  を求め、初回の情緒  $E_1 = \begin{pmatrix} Fr \\ Ar \end{pmatrix}$  を決定する。次に、気分  $M$  を決定する。初回は  $M_1 = E_1$  とする。2 回目以降は 1 回目と同様に  $Fr$  と  $Ar$  を求め、以下の式により情緒  $E_n$  を更新する。

$$E_n = (1-t) \begin{pmatrix} Fr \\ Ar \end{pmatrix} + tM_{n-1} \quad (n \geq 2)$$

$t$  はどれだけ現在の気分を反映させるかの定数である。気分の更新は生じた情緒の種類によって以下の二つに分けられる。

1. 生じた情緒の種類が直前の情緒の種類と異

なる場合： $M_n = E_n$

2.生じた情緒の種類が直前の情緒の種類と同じ場合：

$$M_n = \frac{\alpha}{2}(M_{n-1} + E_n) \quad (\alpha: \text{気分の減衰率})$$

以上により情緒および気分が更新される。

生じた情緒はシステムの表情として出力される。また、現在の気分に応じ、いくつかの応答に対してその応答を選択する確率を変化させる。

### 2.3. 応答の選択

現在の対話の状態をタスクとフェーズの組として表現し、それらに加えシステムの気分、発話行為タイプ、キーワードが含まれているか、ユーザの沈黙時間等に基づいて複数の応答文の候補のスコアリングを行いスコアが最大の応答候補を選択する。ここでタスクとは、雑談・クイズ中・石取ゲームのいずれかである。一方、フェーズとはタスクが行われている・中断している・もう一度行うか現在確認中のいずれかである。

スコアの値は現在の状態に対しユーザの沈黙時間や相手を褒める内容や貶す内容が発話に含まれているか等を報酬とした強化学習により調整される。

最後に応答文を音声合成して出力する。

## 3. システムの実装及び評価実験

### 3.1. ログデータの収集および学習

大学生に子どもを演じてもらい、システム役の人間と実際の対話を想定した模擬対話を行い9人から合計563発話を収集した。これをもとにLIBLINEARおよびCRFsuiteの学習を行った。それぞれで使用した素性を表2、表3に示す。

表2: 発話行為タイプ推定の素性

単語 1, 2, 3gram <sup>*1</sup>
品詞 1, 2, 3gram <sup>*1</sup>
発話文の形態素数
直前のシステムの発話が焦点を含むか

表3: 焦点解析の素性

単語 1, 2gram <sup>*1</sup>
品詞 1, 2, 3gram <sup>*1</sup>
発話文の形態素数
形態素の文中での位置
直前のシステムの発話の発話行為タイプ

<sup>\*1</sup>直前のユーザとその前のシステムの発話の n-gram

### 3.2. プロトタイプシステムの実装

本研究のシステムのプロトタイプを作成した。このプロトタイプに用いたキャラクターの喜び、怒り、悲しみ、平常時での発話状態のグラフィックを図2に示す。

また、今回は音声認識を用いずユーザが話した音声をおペレータが入力する形で対話を行った。

クイズは人手で23問作成した。



喜ぶ 怒り 悲しみ 平常(発話時)

図2: キャラクターの各感情のグラフィック

### 3.3. 評価実験

前述したプロトタイプを用いて小学2年生3人、5年生2人と対話を行った。対話では感情を表出するシステム、しないシステムそれぞれと各7分間対話を行い、対話するシステムの順番はランダムに入れ替えた。合計244発話を収集した。

3.1節で述べたLIBLINEARとCRFsuiteのモデルを用いて収集した発話の発話行為タイプと焦点の推定を行ったところ、それぞれ72.5%、90.8%の精度が得られた。

また、対話後に各システムについて4段階で評価してもらった。評価内容および各システムの平均値を表4に示す。

表4: 各システムのアンケート評価結果

アンケート項目	感情あり	感情なし
上手に会話できていたか	2.8	3.0
話していて面白かったか	3.0	3.0
また話してみたいか	3.0	3.2
仲良くなれそうだったか	3.0	3.0

アンケートの結果から、感情を表情として表出するシステムとしないシステムで、あまり差異は見られなかった。

## 4. まとめ

本研究では子どもを対象とする雑談対話システムの提案を行い、実際にプロトタイプを構築し、発話理解の精度および感情の表出についての評価を行った。

今後、クイズの自動生成機能の実装や、強化学習による応答選択の最適化に関する評価実験を行う。

### 参考文献

[1] 東中竜一郎, “雑談対話システムに向けた取り組み”, 人工知能学会第70回言語・音声理解と対話処理研究会(2014)

[2] 西村竜一他, “実環境研究プラットフォームとしての音声情報案内システムの運用”, 電子情報通信学会論文誌 J87-D-II(3) pp.789-798(2004)

[3] <https://www.csie.ntu.edu.tw/~cjlin/liblinear/>

[4] <http://www.chokkan.org/software/crfsuite/>

[5] J.A.Russell, “A circumplex model of affect”, Journal of Personality and Social Psychology, 39(6), pp.1161-1178(1980)

[6] 高村大也他, “スピンモデルによる単語の感情極性抽出”, 情報処理学会論文誌, 47(2) pp.627-637(2006)