

## 感情判断に基づくコンピュータ会話応答文生成方式

## A Computer Conversation Reply Sentence Generation Method Based on Emotion Judgment

岡崎 雄亮† 渡部 広一† 河岡 司†  
Yusuke Okazaki Hirokazu Watabe Tsukasa Kawaoka

## 1. はじめに

近年、コンピュータは急速に発展している。そして、現在ではコンピュータは様々な場面で必要不可欠なものである。それに伴い、コンピュータを扱う側にも技術と知識が要求される。そのため、誰もが簡単に使えるというわけではない。そこで、誰もが使いやすい、人間にとって優しいコンピュータが必要である。

人間にとって優しいコンピュータには、コンピュータがより人間に近づき、人間とコンピュータとのコミュニケーションが必要であると考えられる。コミュニケーションをとるためには、「会話」という要素が欠かせない。会話において人間同士ならば、相手の感情を読み取るのが普通である。これと同様に、人間とコンピュータとの会話でも感情を判断することが必要である。そして、会話を行うためには、感情を判断した上でコンピュータからの応答が必要になる。そこで本稿では、円滑な会話実現の第一歩として、感情判断に基づく応答文生成方式を提案する。

## 2. 関連技術

## 2.1 概念ベース

概念ベース<sup>[1]</sup>とは語(概念)と意味(属性)のセットを約9万語蓄積してあるものである。この概念と属性のセットには、その重要性をあらわす重みが付与される。ある「概念A」は、その概念と関連が深く、その概念の意味となると考えられる「属性 $a_i$ 」と、その属性の重要性を表す「 $w_i$ 」の組の集合で表される。概念Aは以下のように表される。

$$A = \{(a_1, w_1), (a_2, w_2), \dots, (a_n, w_n)\}$$

属性 $a_i$ を概念Aの一次属性と呼ぶ。一次属性 $a_i$ を一つ概念と見なせば、 $a_i$ からさらにその一次属性を導くことができ、 $a_i$ の属性 $a_{ij}$ を概念Aの二次属性という。これを展開していくと、一つの概念Aはn次属性まで持つことができる。概念ベースの例を表1に示す。

表1: 概念ベースの例

| 概念  | (属性, 重み)  |           |     |
|-----|-----------|-----------|-----|
| 自動車 | (車, 0.9)  | (軸, 0.45) | ... |
| 太陽  | (日光, 0.2) | (月, 0.05) | ... |
| :   | :         | :         | :   |

## 2.2 関連度計算

概念Aと概念Bの関係の深さを定量的に表すのが関連度計算<sup>[2]</sup>という方法である。関連度は、0以上1以下の連続的な数で表され、概念同士の関連が大きいほど関連度は高くなる。この関連度を求める計算は、それぞれの概念を二次属性まで展開し、その重みを利用した計算によって最適

な一時属性の組み合わせを求め、それらが一致する属性の個数を評価することで算出する。関連度計算の例を表2に示す。

表2: 関連度の例

| 基準概念 | 対象概念 | 関連度   |
|------|------|-------|
| 飛行機  | 航空機  | 0.418 |
|      | 自動車  | 0.057 |
|      | 花    | 0.003 |

## 2.3 意味理解システム

意味理解システム<sup>[3]</sup>とは単文入力に対して6W1H(who, what, when, where, whom, why, how)と用言という8個の成分に分けることの出来るシステムである。分けられた語は表3のフレーム(意味フレーム)に分割される。本稿では、この中からwho, what, when, where, whom, 用言の6つのフレームを用いることとする。

表3: 意味フレーム

| who | what | when | where | whom | why | how | 用言 |
|-----|------|------|-------|------|-----|-----|----|
|     |      |      |       |      |     |     |    |

## 2.4 感覚判断システム

感覚判断システム<sup>[4]</sup>は、ある語(名詞)に対して人間が常識的に抱く印象を形容詞・形容動詞の形で出力するシステムである。『痛い』『臭い』といった人間が五感で感じる印象を“五感感覚語”、『めでたい』『不幸な』といった五感以外で感じる印象を“知覚語”と呼び、この2つを総称して“感覚語”と呼ぶ。感覚判断システムの例を表4に示す。

表4: 感覚判断システムの例

| 入力語 | 感覚語         |
|-----|-------------|
| レモン | 黄色い, 酸い     |
| 空   | 広い, 青い, 大きい |

## 2.5 感情判断システム

感情判断システム<sup>[5]</sup>とは入力された語から感情を判断するものである。感情判断システムでは、“感情語”を扱う。感情語とは発言の時点に瞬間的に発生すると考えられる『喜び』『安心』『悲しみ』『恐れ』『怒り』『落胆』『後悔』『恥』『罪悪感』に加え、感情『なし』の計10種である。結果はこの10種類のうちの1つ、または複数となる。感情判断システムの例を表5に示す。

表5: 感情判断システムの例

| 入力文      | 感情          |
|----------|-------------|
| 私は宝石を買った | 喜び          |
| 私は事故に遭った | 後悔, 落胆, 悲しみ |

† 同志社大学大学院工学研究科

Graduate School of Engineering Doshisha University

### 3. 応答文生成システム

#### 3.1 システムの概要

本研究では、感情判断に基づいた応答文生成システムを構築した。図1にその流れを示す。

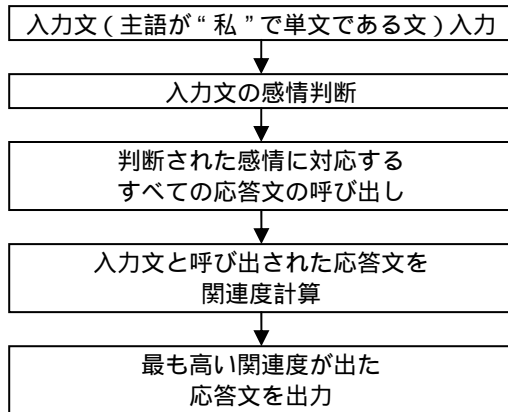


図1: システムの流れ

このシステムでは、主語を“私”に限定することによって、話者自身の感情に見合った応答文を出力することができる。また、入力文に特定の語があれば、特定の応答文を出力するといったような一対一の対応にするのではなく、一つの入力文に対して複数の応答文の候補がある中で、その入力文に最もふさわしい応答文を出力することができる。例えば、“私は結婚した”という入力文に対して、応答文の候補は“おめでとう”“よかったですね”“嬉しかったですよ”等があるのだが、この場合は最もふさわしい“おめでとう”が出力される。つまり、入力パターンによらない応答システムである。これは、概念ベース・関連度計算を用いることで、入力文と応答文の関連性を調べ、結果的に最も関連が高いものを選択できるからである。

#### 3.2 応答文データベース

応答文の候補を格納した応答文データベースを作成した。この応答文データベースには、感情判断で出力される全9種類の感情に対応するように、44文の応答文を格納した。更に、その応答文を一語で特徴付ける単語(以降、簡略形と呼ぶ)を共に格納した。例として、“おめでとう”は“めでたい”となる。これは入力文との関連の妥当性を定量的に評価する関連度計算をより正確に行うための単語である。応答文を呼び出しする際は、判断された感情に対応するすべての簡略形を呼び出し、入力文との関連度計算を行う。表6に応答文データベースを示す。

表6: 応答文データベース

| 感情  | 応答文      | 簡略形  |
|-----|----------|------|
| 喜び  | おめでとう    | めでたい |
| 喜び  | 楽しかったですよ | 楽しい  |
| :   | :        | :    |
| 悲しみ | 悲しかったですよ | 悲しい  |
| :   | :        | :    |
| 怒り  | 腹立たしいですね | 立腹   |
| :   | :        | :    |

#### 3.3 応答文選択

本節では、入力文入力から応答文選択までのシステム全体の流れを具体的に例を用いて示す。なお、例は“私は試合に優勝する”という文を用いる。

まず、意味理解システムを用いて入力文を who, what, when, where, whom, 用言の6つのフレームに分ける。この時、サ変名詞は“する”を除いて名詞形にする。本節の例の場合は表7のようなになる。次に、入力文を感情判断システムに掛ける。ここでの感情は『喜び』となる。次に、判断された感情に対応した全ての応答文の簡略形を呼び出し、その一次属性を取得(図2)する。そして、取得した全ての一次属性とフレーム分けされた who フレーム以外の語との関連度を取る(図3)。最後に、最も高い関連度が出た語が一次属性に含まれる応答文を選択する。ここでは、“めでたい”の一次属性“勝つ”と“優勝”の関連度が最も高いため、応答文は“おめでとう”となる。

表7: 入力文のフレーム分け

| who | what | when | where | whom | 用言 |
|-----|------|------|-------|------|----|
| 私   | 試合   |      |       |      | 優勝 |

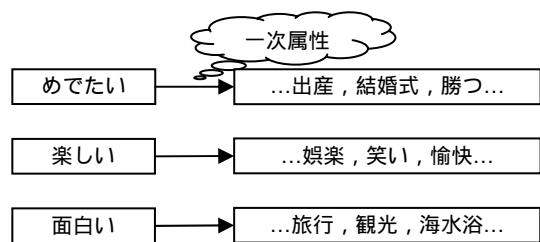


図2: 一次属性の取得

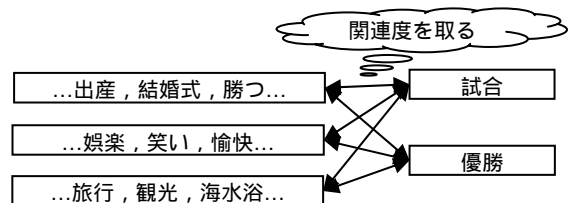


図3: 関連度計算

### 4. 評価

感情『なし』を除く全9種類の感情について、それぞれその感情を表す文章を入力文にする。その際制約があり、以下の通りである。

- 必ず“私は”から始まる文章であること
- 単文であること

また、感情判断システムにおいて正しい判断結果を出力した入力文を対象とし、本システムの有効性を調査した。その入力文の内訳は『喜び』が34個、『悲しみ』が57個、『落胆』が62個、『罪悪感』が12個、『後悔』が50個、『怒り』が14個、『恐れ』が10個、『恥』が9個、『安心』が4個である。判定は目視で評価し、入力文に対して矛盾を感じない応答文ならば正解、入力文に対して明らかに間違った応答文ならば不正解とした。全ての評価文に対する正解の割合を精度とし、結果を図4に示す。

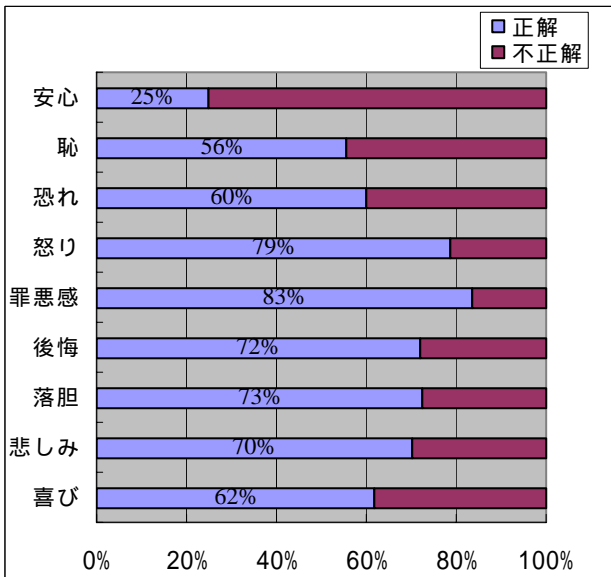


図4：実験結果

5. 考察

表8に成功例を示す。

表8：システムの成功例

| 入力文      | 感情  | 応答文      |
|----------|-----|----------|
| 私は内定を貰った | 喜び  | おめでとう    |
| 私は物を盗んだ  | 罪悪感 | 君には失望したよ |
| 私は病気になった | 悲しみ | 苦しかったですよ |
| ：        | ：   | ：        |

表9に失敗例を示す。

表9：システムの失敗例

| 入力文        | 感情  | 応答文       |
|------------|-----|-----------|
| 私は病気が治った   | 喜び  | 楽しかったですよ  |
| 私は試合に勝った   | 喜び  | 美味しかったですよ |
| 私は息子を誘拐された | 悲しみ | 不味かったですよ  |
| ：          | ：   | ：         |

上記の失敗例の考察と問題点を示す。まず、“私は病気が治った”という入力文であるが、これは“楽しい”の一次属性に“不快”があり、この“不快”と“病気”との間で最も高い関連度が出たために“楽しかったですよ”という応答文になった。また、“私は息子を誘拐された”という入力文であるが、これは“不味い”の一次属性に“男性”があり、この“男性”と“息子”との間で最も高い関連度が出たために“不味かったですよ”という応答文になった。このように、一次属性の中に不適切なものが含まれるという問題点が見つかった。

6. システムの改良と再評価

6.1 システムの改良

5.2節で示した問題点を解決するためには、一次属性そのものを見直す必要があると考えた。そこで、概念ベースが

ら獲得していたすべての属性を見直し、目視で明らかに不適切なものだけを省き、残ったものを新たにデータベースに格納し直した。これを属性データベースと呼ぶ。そして、属性を獲得する際は属性データベースから呼び出すことにした。表10に例を示す。

表10：属性データベース

| 感情  | 簡略形  | 属性                |
|-----|------|-------------------|
| 喜び  | めでたい | 勝つ, 順調な, 晴れやかな... |
| ：   | ：    | ：                 |
| 悲しみ | 苦しい  | 病, 苦痛, 重労働...     |
| ：   | ：    | ：                 |
| 恥   | 照れる  | 照れる, 赤らめる...      |
| ：   | ：    | ：                 |

6.2 再評価

改良したシステムの評価を行った。評価方法は4章の通りであり、入力文の個数も同じである。図5に結果を示す。また、図6にシステム改良前と改良後のシステム全体の精度を示す。

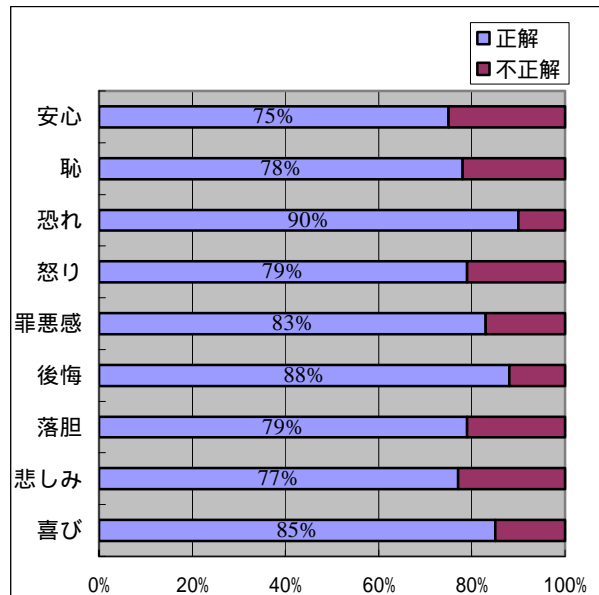


図5：システム改良後の実験結果

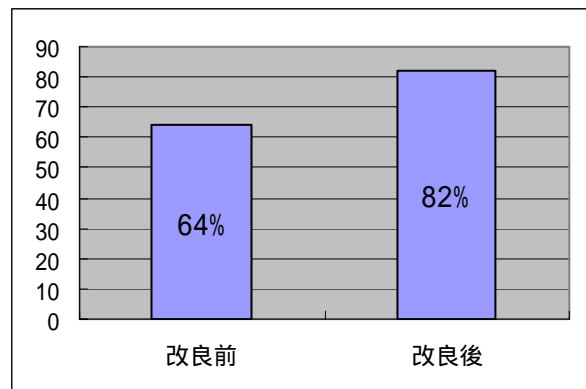


図6：システム全体の精度

図6の様に、システム全体として18%の精度向上を挙げることが出来た。

## 7. おわりに

人間とコンピュータとの円滑な会話実現を目指し、コンピュータの感情判断に基づいた応答文生成システムを構築した。具体的には、応答文データベースの作成、入力文と応答文の関連の妥当性を定量的に評価する関連度計算方法の考案を行った。また、システム作成後に見つかった一次属性の中に不適切なものが含まれるという問題点を解決するため、属性データベースの作成を行った。これによりシステム全体として18%の精度向上を挙げることが出来、コンピュータからの応答が可能になった。その結果、感情を踏まえた人間とコンピュータとの円滑な会話実現の第一歩が踏み出せたのではないかと考える。

## 謝辞

本研究は文部科学省からの補助を受けた同志社大学の学術フロンティア研究プロジェクトにおける研究の一環として行った。

## 参考文献

- [1] 広瀬幹規, 渡部広一, 河岡司, “概念間ルールと属性としての出現頻度を考慮した概念ベースの自動精練手法”, 信学技報, TL2001-49, pp.109-116, 2002
- [2] 渡部広一, 河岡司, “常識的判断のための概念間の関連度評価モデル”, 自然言語処理, Vol.8, No.2, pp.39-54, 2001
- [3] 篠原宜道, 渡部広一, 河岡司, “常識判断に基づく会話意味理解方式”, 言語処理学会第8回年次大会発表論文集, B6-2, pp.651-654, 2002
- [4] 渡部広一, 堀口敦史, 河岡司, “常識的感覚判断システムにおける名詞からの感覚想起手法”, 人工知能学会論文誌, Vol.19, No.2, pp.73-82, 2004
- [5] 風間勇志, 植野義孝, 渡部広一, 河岡司, “常識的感情判断と主体語処理”, 情報科学技術フォーラム FIT2002, E-28, pp.137-138, 2002