

対面コミュニケーションと SNS とを行き来するエージェント

An agent which moves to face-to-face communication and SNS

藤倉 将平*¹
Fujikura Shohei小川 義人*¹
Ogawa Yoshito菊池 英明*²
Kikuchi Hideaki*¹ 早稲田大学 人間科学研究科
Graduate School of Human Sciences, Waseda University*² 早稲田大学 人間科学学術院
Faculty of Human Sciences, Waseda University

In this challenge, we aimed at the realization of an agent system which can make communication in use of the information of user's "personality" on the 2 media, face-to-face dialogue and Twitter. The system automatically estimates the personality from user's utterances or tweets on each medium, then replies an appropriate response for it. We constructed the agent dialogue system and confirmed the effect of the personality estimation.

1. はじめに

1.1 本研究の目的

音声対話を通じて人間の特定のタスクを請け負う新しいアシスタントサービスとして Apple の Siri や NTT docomo のコンシェルジュサービスなどが台頭し、注目を集めている。いずれも特徴的であるのは、人を模したエージェントを用意することでユーザが難しく面倒な操作を行う必要性を減らしつつ、ユーザのインタラクション欲求を高めていることである。

従来研究では、システムのデザインを工夫したり、話題を増やして多様な対話シーンに対応することでインタラクションの新規性を高める試みがなされてきたが、継続性の評価は十分に行われていなかった。インタラクションの新規性を高めるデザインは短期的にはユーザの興味を引くものの、長期的な利用においては飽きられてしまう可能性がある。また、ユーザが対話を続けたいと感じる話題の要因が明確になっていないため、話題を増やすには人手によるか、大量のデータからの場当たりの生成によるしかなく、長期的に新規性を維持することが難しい。

そこで我々は、人同士のインタラクションを参考に、毎回の対話において「次回も続けたい」と感じる要因を分析することで、対話継続欲求の高いシステムのデザイン方法論の確立を目指してきた。第一に人同士の対話を円滑化させる「社会性の高い対話」に着目し、会話分析によって要因を明確化した。第二に相手の関心を引く「ユーモア表現を含む会話」に着目し、ユーモア発話を行う対話システムを構築して評価実験を行った。その結果、人同士の対話において相手の発話行動を限定しない事、相手の話を聞いている実感を与える事、システム発話において意外性の高いユーモアを使う事が有効であった[宮澤 12]。

本研究では対話システムがユーザのパーソナリティを自動的に推定し、その情報を活用した応答を返すことによってユーザが対話を続けたいと思えるシステムを構築する。このシステムでは全てのユーザに対して画一的な反応を返すのではなく、それぞれのユーザのパーソナリティごとに反応を変えるというアプローチをとる。このシステムはユーザが過去に発言したテキストデータからその人が

どのようなパーソナリティをどの程度持っているのか数値化することにより、ユーザのパーソナリティを自動的に推定する。

さらに、本研究は一つのエージェントが同じユーザに対して対面(音声対話)と Twitter の両方で双方のメディアの特性に応じたコミュニケーションを行えるようにすることを目指している。ユーザは音声対話におけるパーソナリティと Twitter 上でのパーソナリティの 2 つを持つと仮定する。本稿では、2 つのメディアにおけるパーソナリティ自動推定の精度を確認し、推定したパーソナリティ情報を用いることによる対話システムとの対話継続欲求の向上についてメディア毎に比較する。

1.2 対話例

本研究において実現させる対話例を以下に記す。

S:対話システム	U:ユーザ
U:やばい	
	(もしユーザが悩みがちな人ならば)
S:どうしたの?大丈夫?	
	(もしユーザが怠惰な人ならば)
S:なんだ遅刻したのか	
	(もしユーザが好奇心の強い人ならば)
S:なにかすごい物でも見つけた?	

ユーザのパーソナリティを BigFive 尺度[斉藤 01]によって表現する。ユーザのパーソナリティは外向性、神経症傾向、開放性、調和性、誠実性をそれぞれ 0~50 の点数で表される。また BigFive 尺度のそれぞれに対して正の因子負荷を持つ項目群をポジティブ、負の因子負荷を持つ項目群をネガティブと呼ぶ。

2. ユーザのパーソナリティ自動推定

2.1 推定方法

ユーザの発言それぞれを BigFive 尺度に則り分類することで、そのユーザがどのようなパーソナリティを持っているのか推定する。BigFive 尺度のそれぞれにおいてポジティブとネガティブとで二値分類を行う。分類の結果としてポジティブなものと判断された場合は1、ネガティブなものと判断された場合は0を返すものとする。それぞれの尺度において分類された直近の 50 個までのデータを足し

合わせた 0~50 の得点をパーソナリティ得点とし、そのユーザの持つパーソナリティとする。それぞれの尺度においてポジティブな文とネガティブな文がそれぞれ等確率でユーザの発言に出現すると仮定すると $p=0.5$ の二項分布となる。またこの2項分布を正規分布に近似する。正規分布において平均から 3σ に分類の履歴の長さが位置するとして σ 値を計算する。それぞれの尺度において点数が平均から 1σ 以上離れていた場合はそのユーザのその尺度のパーソナリティを持つと推定される。例えば外向性において平均より 1σ 以上得点が大きかった場合は外交的であると推定され、平均より 1σ 以上得点が小さかった場合は内向的であると推定される。

2.2 分類器の構築

半教師多項式ナイーブベイズ分類器[Jiang 11]を使う。ラベル有り学習データセットは後述する調査を行う事によって収集した。ラベルなしデータセットはツイート[Twitter]と呼ばれるテキスト集合を用いた。素性として形態素を扱うため文を形態素解析器である MeCab に渡すことで形態素に変換する。

2.3 ラベル付きデータセット収集のための調査1

ラベル付きデータセットを収集するため、自由回答のアンケート調査方式をとった。被験者は、あるパーソナリティを保持している人がどのような発言を行うか意識して文を作成した。被験者は10名となった。

2.4 ラベル付きデータセット収集のための調査2

この実験ではアンケート回答者の意識に登らないような発言表現を収集することを目的とした。被験者に対して会話コーパスの書起しと Twitter から集めた「ツイート」をランダムに提示し、被験者はその文から発言者がどのようなパーソナリティをもっているか判断し、ラベル付けを行う。被験者は10名となった。

2.5 分類器の評価実験

表1 分類器の評価結果

BigFive 尺度	ラベル	分類失敗文数	総文数	分類成功率
外向性	ポジティブ	136	464	70.69%
	ネガティブ	81	210	61.43%
神経症傾向	ポジティブ	90	317	71.61%
	ネガティブ	82	262	68.70%
開放性	ポジティブ	98	248	60.48%
	ネガティブ	75	230	67.39%
調和性	ポジティブ	57	208	72.60%
	ネガティブ	74	199	62.81%
誠実性	ポジティブ	77	206	62.62%
	ネガティブ	117	366	68.03%

表1は分類器の交差検定の結果を表している。フォールド数は5とした。分類成功率は10項目を平均して約66.66%となった。

3. 対話システム

3.1 対話システムの構成

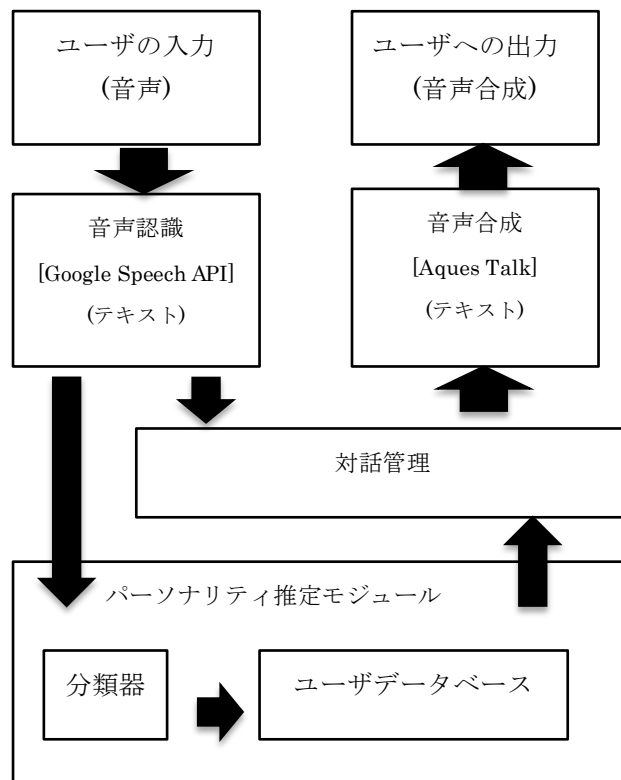


図2 音声対話における対話システムの構成

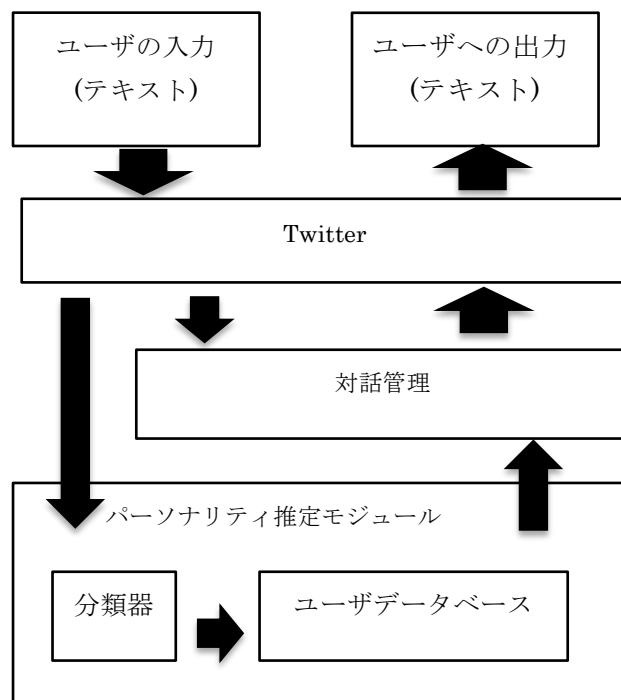


図3 Twitterにおける対話システムの構成

図 2 は[河原 06]を参考にして構成された音声対話におけるシステムを表している。音声認識[Google Speech API]によって音声をテキストに変換し対話管理部とパーソナリティ推定モジュールへ渡す。対話管理部ではユーザデータベースに保存されているパーソナリティ情報を参照し 2.1 で述べたパーソナリティ推定方法に則ることでユーザのパーソナリティを推定し、応答文を決定する。対話管理部によって決められた文を音声合成[Aques Talk]することで、音声としてユーザに応答を返す。

図 3 は Twitter における対話システムを表している。ユーザが Twitter にシステム宛てのツイートを投稿し、システムはユーザ宛てのツイートを行うことで対話を行う。対話管理部では音声対話と同じく、ユーザのパーソナリティ情報を参照する事によって応答を決定する。システムのユーザへの応答も Twitter に投稿される。

パーソナリティ推定モジュールは別のスレッドにおいてユーザの発言を受け、それぞれの尺度において極性判別した結果をデータベースに保存する。

3.2 用意したパーソナリティ情報を活用した発話表現

本研究において実現させる対話例を以下に記す。

用意した対話表現の一部

S:対話システム U:ユーザ

U:やばい

(神経症傾向においてネガティブと推定されたら)

S:どうしたの?大丈夫?

(誠実性においてネガティブと推定されたら)

S:なんだ遅刻したのか

(開放性においてポジティブと推定されたら)

S:なにかすごい物でも見つけた?

U:だるい

(誠実性においてネガティブと推定されたら)

S:怠けてないで早く学校行こう!

(外向性においてネガティブと推定されたら)

S:たまには外に出てみたらどう?

(外向性においてポジティブと推定されたら)

S:少しは家でのんびりと休んだら?

4. 質問用紙法とパーソナリティ推定システムにおける比較

従来の BigFive 尺度でパーソナリティを測定するために使用されてきた質問用紙によるユーザのパーソナリティと、本研究で使用しているパーソナリティ推定システムによって推定されたユーザのパーソナリティを比較する。質問用紙による質問内容は[堀 01]を参考にした。被験者は大学生被験者 11 名となった。

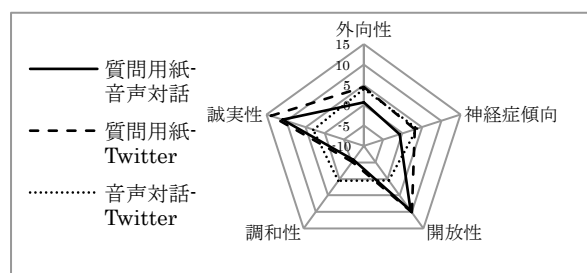


図 4 それぞれの手法における全被験者のパーソナリティの得点の差の平均

図 4 はそれぞれの手法における全被験者のパーソナリティの得点の差の平均を表している。例えば図の「質問用紙-音声対話」のグラフは質問用紙によって測定されたパーソナリティ得点から音声対話においてシステムに推定されたパーソナリティ得点の差の、全被験者における平均を示したものである。「質問用紙-音声対話」のグラフ、「質問用紙-Twitter」のグラフから調和性は、質問用紙では低く得点されていても、音声対話と Twitter におけるパーソナリティ推定システムによると高い得点を持つと推定されている傾向が強かったことが確認出来る。開放性と誠実性においては、パーソナリティ推定システムによる推定結果は質問用紙の得点より低い傾向が強かった。これは対話を行う際には、「無駄」、「めんどい」のような開放性、誠実性においてネガティブに捉えられる表現が多かったこと、また調和性において「〜ください」のようなポジティブに捉えられる表現が多かったことが考えられる。「音声対話-Twitter」のグラフから、外向性において音声対話で推定された得点の方が Twitter で推定された得点の方よりも少し高い傾向にあることが確認出来る。これは音声対話において「〜に行きたい」などの外向性の高くなるような表現が多かったことが考えられる。また開放性、誠実性については少し低い傾向にあることが確認出来る。これはまた Twitter 上において「無駄」、「めんどい」のような開放性、誠実性の得点が低くなるような表現が多かった事が考えられる。

5. 対話システムの評価

5.1 目的

構築した対話システムについて、ユーザのパーソナリティを自動的に推定し、その情報を活用した応答を返すことによってユーザが対話を継続したいと思えるようになるか評価する。

5.2 評価実験概要

パーソナリティ情報を用いた発話と用いない発話の2つの対話をユーザとシステムの間で行う事によって比較し、検証を行った。最初に被験者はシステムによるパーソナリティ推定のために 50 回対話を行う。次にパーソナリティ情報を扱う対話、もしくは扱わない対話のどちらかを先に 30 回行いアンケートに答える。最後に残りの対話を同じく 30 回行い、アンケートに答える。被験者は 11 名となった。

5.3 調査内容

調査内容として以下の項目を用意した。

1. このシステムは自分のことをよくわかっている [1~7段評価]
2. (1の設問に対して) なぜそう感じたのか
3. 今後もこのシステムと対話をしたい [0~6 評価]
4. システムとの対話をもっと長く続けたい [0~6 評価]
5. (3と4の設問に対して) なぜそう感じたのか

5.4 調査結果

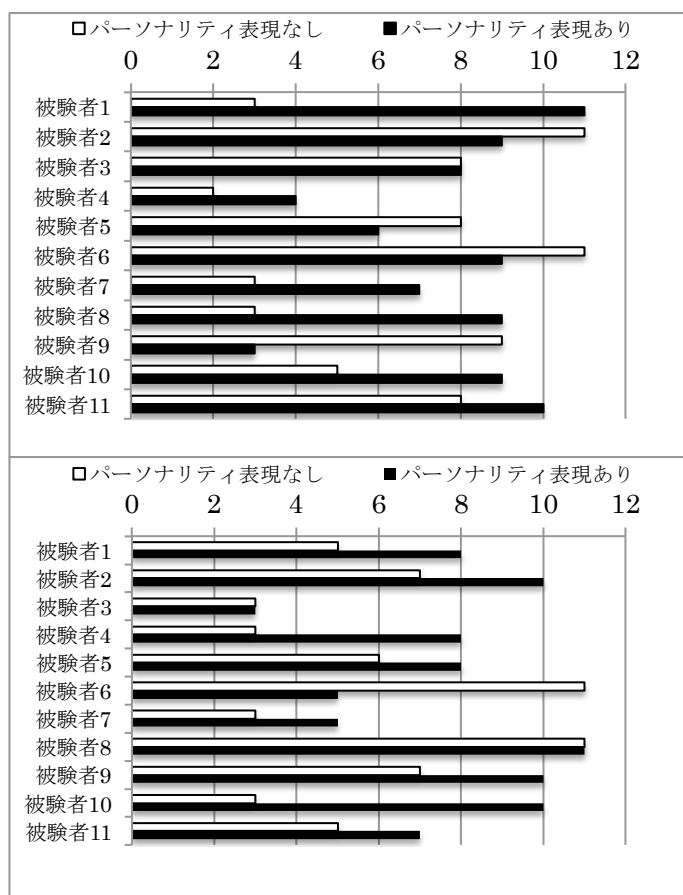


図5 対話継続欲求の評価結果
(上図は音声対話, 下図は Twitter)

図5はパーソナリティ情報を扱わない対話と扱う対話で継続欲求がどのように変化するか表している。調査内容の3, 4番目の項目を加算した点数を対話継続欲求としている。音声対話によるパーソナリティ情報を扱わない対話の対話継続欲求の平均値は7.1, パーソナリティ情報を扱う対話の平均値は8.5, Twitter上でのパーソナリティ情報を扱わない対話の平均値は6.4, パーソナリティ情報を扱う対話の平均値は8.5となった。音声対話とTwitterでの対話ともに, 対応のあるt検定によりパーソナリティ情報を扱わない対話とパーソナリティ情報を扱う対話の間で有意差は見られなかった。音声対話においてパーソナリティ情報を扱う対話によって対話継続欲求が向上した被験者は6名, Twitterでの対話においてパーソナリティ情報を扱う対話によって対話継続欲求が向上した被験者は8名となった。

6. まとめ

ユーザのパーソナリティを自動的に推定し, その情報を活用した応答を返す事によって音声対話とTwitterでの対話において共に過半数の被験者の対話継続欲求の向上を確認する事ができた。質問用紙によって測定されるパーソナリティと, 音声対話やSNSにおける発話行動から推定されるパーソナリティの間には差が生じること, また音声対話とSNSでは推定されるパーソナリティが異なることが確認出来た。これは相手や場に応じた発話行動を行った結果と考えられる。

本研究の今後の課題を以下に述べる。2.5で示した分類器の性能が約66.66%に留まったが, 2.4で集めたラベル付きデータセットの中にはポジティブとネガティブのどちらともつかないような表現が含まれていた。これらの妥当性を検証することで分類の成功率を高める事ができると考えられる。また3.3で用意したパーソナリティ情報を活用した発話表現については, これらが妥当な表現であるのか調査する必要があると考えられる。本研究では音声対話とTwitterにおいて, メディアごとにパーソナリティを推定し応答を選択することで対話継続欲求の向上を狙ったが, 今後は2つのメディアを行き来することによりどちらかのメディアからのみでは得る事が出来ないような情報を得ることを検討していく。

参考文献

- [宮澤 12] 宮澤幸希, 常世徹, 榎井祐介, 松尾智信, 菊池英明: "音声対話システムにおける継続欲求の高いインタラクションの要因", 電子情報通信学会論文誌 A, Vol. J95-A, No.1, pp.27-36, 2012.
- [齊藤 01] 齊藤崇子, 中村知靖, 遠藤利彦, 横山まどか: "性格特性用語を用いた BigFive 尺度の標準化", 九州大学心理学研究 || 2 ||, p135-144 (2001).
- [Jiang 11] Jiang Su, Jelber Sayyad Shirab, Stan Matwin. : "Large Scale Text Classification using Semi-supervised Multinomial Naive Bayes", ICML, USA, 2011.
- [Twitter] Twitter : (<http://twitter.com>).
- [河原 06] 河原達也 荒木雅弘: "知の科学 音声対話システム", 人工知能学会(編), オーム社, 東京, 2006.
- [Google Speech API] Google : (<http://www.google.com/intl/en/chrome/developer/speech.html>).
- [Aques Talk] Aquest : (<http://www.aquest.com/products/aquestalk.html>).
- [堀 01] 堀洋道: "心理尺度集 I", 山本眞理子(編), サイエンス社, 東京, 2001.