

共起・共助マッチングに基づく対話支援ロボットの開発に向けた人同士の会話解析

An analysis of human dialogue based on co-occurrence or mutual assistance relation for a developing chat robot

山口 亨*¹ 下川原 英理*¹ 五味 怜央奈*¹ 相澤 秀和*¹ 岩崎 真也*¹
Toru Yamaguchi Eri Sato-Shimokawara Reona Gomi Hidekazu Aiazawa Shinya Iwasaki

*¹首都大学東京 システムデザイン研究科
Tokyo Metropolitan University, Graduate School of System Design

Authors have been focusing on the Community-centric system which ties individual persons each other. Previous our study proposed a matching method for co-occurrence and mutual assistance relation using personal attributes. Matched persons might have a chance to make a relation. Then we are developing a robot which supports to create a good relation or communities. Therefore, this research focused on an analysis of human conversation finding an evaluation of human relation. This paper reports the result of the analysis using a heart rate or sound features. The results indicate the future tasks and directions.

1. はじめに

Robi や Pepper などコミュニケーションロボットが身近なところで存在し、簡単な案内をしたり、子どもたちを楽しませたりしている。2020年の東京オリンピック・パラリンピックでもこのようなロボットの活躍が期待されている。しかしながら、現状のコミュニケーションロボットでは、人間のような「気の利いた」「きめ細かい」案内は難しい。そこで、我々が提案しているのは、ロボットが窓口となり、問い合わせ分野や言語などを推定し、それらを得意とする人とマッチングすることである。さらに、案内スタッフも、それぞれ、英語が得意、歴史に詳しい、など得意分野がある。この得意分野を活かせるようなマッチングが重要であると考えている。図1のStep 1に示すように、位置情報や心拍数、ロボットとのインタラクション履歴などから、ユーザーの特徴を抽出する。人による観光案内が、今までに行ったことのない場所を案内するだとか、行ったことはあるが季節よってことなるイベントが開催されているなど新しい発見を提供するような案内を行うことが可能なのは、プロフェッショナルな案内スタッフが、うまくユーザーのニーズを聞き出していたからであろう。その点で、第三次人工知能ブームとも呼ばれる現在、ビッグデータなどの大量データを活用したディープラーニングが有効であろうと考えている。図1のStep 2に示すように、個人情報秘匿プライバシーに配慮してデータを収集することで、ビッグデータ活用につなげる。大量のデータから見つけ出された「ここに行ったことのある人は、ここに行くといい」といった推論に結びつけ、観光客と案内スタッフとの橋渡し役になるロボットの知識として利用するのである(図1のStep 3)。我々はこのソーシャルロボティクスと人工知能を活かしたコミュニティ支援をCcAI (Community-centric Artificial Intelligence) と呼んでいる。ロボットを活用すると、Face to Faceの案内は勿論のこと、テレプレゼンスロボットとして遠隔にいるスタッフと繋げることも可能となる。これにより、東京周辺の人だけでなく、全国の人材、特に地方の人材を活かした仕組みが可能になると考えている。

ここまで観光を例にとって話を進めてきたが、人と人とを結

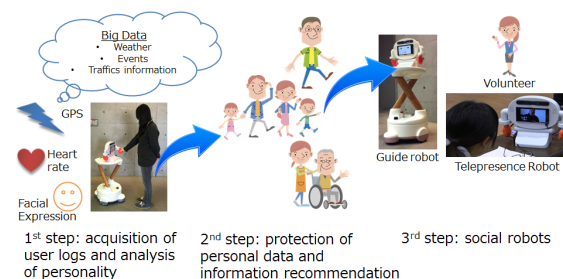


図1: ユーザーのニーズに合わせたソーシャルロボットがつなぐコミュニティ

びつけるというのは、何も案内や受付に限ったことではない。我々はこれまで、総合防災プロジェクト*¹を通して、地域コミュニティの重要性を示唆してきた。ここでもやはり重要なのが、人と人とを結びつけるマッチング技術である。人の興味や関心、能力などから、同じ興味や関心を持つ人同士、お互いに助け合うことができる人同士を見つけ出さなければならない。被災時だけでなく、日常生活から、互いの興味関心や能力(「できること」だけでなく「できないこと」を知ることも重要である)を活かしたコミュニティづくりが必要であると感じている。さらに、それらの人同士が出会ったときに、スムーズに会話できるように支援するシステムが必要である。

そこで、本稿では、人の特徴量を用いて発見された、同じ興味や関心を持つ人同士(本稿ではこの関係を「共起」と呼ぶ)や、互いの得意なことと不得意なことを補う人同士(本稿ではこの関係を「共助」と呼ぶ)を見つけ出し、対話を行った際の生体データの変化に着目し、マッチングの効果について検証するとともに、ロボットの介入について検討する。

2. 共起共助マッチング

前節で述べたように、人と人とを結びつける要素の一つとして、我々は、その人のもつ興味関心や能力に着目している。従ってこれを抽出、獲得、推定する必要がある。興味関心を推定する手法として、Webインテリジェンスでは、購入履歴や閲覧履歴などを利用している[Lee 02]。GPSなどを用いた

*¹ <http://bousaipj.tokyo-sangaku.jp/>

位置情報による興味・関心を推定する研究もおこなわれている [Zheng 09]. ただし単なる頻度ではなく、滞在時間や滞在目的、滞在中の行動なども見る必要であろう。例えば、神田らは、科学館の来館者の移動奇跡を RFID タグを用いてトラッキングすることで、典型的な行動パターンの発見や来館者の人間関係推定などを行っている [神田 08]. 我々は、同様に iBeacon を用いた局所位置情報を用いた、滞在時間などから興味関心を推定する手法について研究を進めている。ポスターセッションを対象に、スマートフォンのセンサを利用して立ち止まりを検出し、そこから興味関心が推定できないかを検証している [本橋 16]. 能力を推定するには、資格や職歴などが役にたつがそれをどのようにして聞き出すかが重要である。本人に聞くことはもちろんだが、他者推薦も有効ではないかと考えている。そこで、ロボットと人との会話から、特徴となるキーワードを抽出し、興味関心の情報を得る手法についても検討している [Sato-Shimokawara 15, Suzuki 16]. ロボットの対話をコントロールすれば、直接的に「好きですか?」「興味はありますか?」と尋ねることもでき、より必要な情報に絞って抽出することができる。このロボットの対話は先に述べた他者推薦にも有効であると考えている。最も古典的な方法はアンケートに記入してもらう方法であろう。アンケートの欠点は頻繁に行うことが難しく、興味関心の移り変わりに対応できないことである。しかし直接的に本人の好みを知ることができる。アンケートやインタビューなどから始め、どのような聞き方をすれば特徴を抽出しやすいのかといった知見を得て、ロボットの対話生成や対話コントロールに応用することも考えている。このように様々な手段を用いて興味関心や能力を抽出することが期待でき、センサーネットワークやマルチモーダルセンシング処理に関する研究と連携して進める必要がある。つまりログデータやビッグデータ解析などの研究とも密接にかかわってくる。本稿では、それに続く対話支援に着目するため、まず記述式のアンケートによって興味関心の抽出を行い、それをもとに対話ペアを見つげ出した。

3. 対話支援

共起共助マッチングなどを経て、よい関係を築けそうなペアを発見しても、その関係を実際に築ききっかけが必要となる。一章で述べたように観光案内であれば「案内」という明確な目的があり会話が始まることになるが、日常的なコミュニティ形成では、まず何から話し始めたらいいのか、挨拶と自己紹介で終わってしまったら関係作りにならない。共起や共助の関係を見出しでも、実際のつながりに活かさなければ意味がない。そこで、我々はロボットが仲介役として対話に参加することを考えている。図 2a 挨拶や自己紹介が終わったところを見計らって、共起の話題や、共助を話題を適宜提示することで、人同士の会話が弾む方向へと導くシステムが必要である。これはもちろん観光案内という場面であっても、普段の趣味や興味、得意なことを踏まえ案内することが可能となり、有効であると考えている。

4. 実験と解析

共起共助マッチングの効果、および対話支援ロボットに向けた検証を行うために、学生を被験者として対話実験を行いデータを収集した。

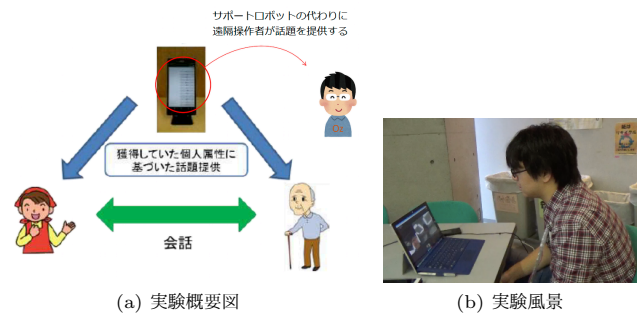


図 2: ビデオチャットによる対話実験

4.1 実験データの収集

本実験では、14名(男性10名, 女性2名)に、興味関心と得意不得意についてそれぞれ30項目のアンケートに5段階評価で回答してもらった。アンケート結果を用いてマッチングを行った。共起はピアソン相関を用いて行い、共助は同じ項目で「得意かつ興味がある」「不得意だが興味はある」人同士とした。マッチング手法については [Gomi 15] に詳しく述べている。今回アンケート項目は30項目を用いているが、項目数については [Sato-Shimokawara 16] においてシミュレーションによって30項目程度であっても十分にマッチング可能であることが示されている。マッチング手法についてはその結果、共起五組 (UserID4 と 6, 4 と 7, 9 と 11, 3 と 12, 11 と 13), 共助五組 (UserID2 と 6, 6 と 13, 9 と 13, 12 と 13, 13 と 14), どちらの関係でもないその他五組 (UserID1 と 11, 10 と 12, 4 と 10, 3 と 5, 3 と 6), の計15組とした [相澤 16]. これらの組み合わせの中には、顔見知りの組み合わせもあれば、ほぼ初対面の組み合わせもある。

この15組を対象として、音声による対話9組、テキストチャットによる対話6組の実験を実施した。

本稿では音声対話についてのみ述べる。テキストチャットの結果については [相澤 16] を参照頂きたい。ペアとなった二人と、サポートロボットを想定した仲介者、の三名で会話を行った。マッチングしたペアの中から、共起、共助、その他それぞれ三組ずつ選び、パソコンで Skype によるビデオチャットで会話を行った (図 2b). 会話中の心拍を取得するために、被験者の耳に心拍センサーを装着した。仲介者は両者の会話を聞ける環境とした。仲介者にはあらかじめ共起または共助の話題を教えておいた。そして仲介者は会話の途切れた時間を計測しており、三秒以上両方の被験者が無言だった場合、テキストチャットを用いて両被験者に話題を一つずつ提示した。被験者には、実験開始とともに自由に会話(フリートーク)をしてもらい、その後仲介役が適宜話題を提示すると説明した。ただし話題を提示するタイミングや提示内容などの詳細な説明はしなかった。

提示する話題は、共起・共助ペアには共起・共助関係に関連する話題を二つと関連しない話題を一つ、その他のペアには無作為に選んだ話題を三つの話題を提示した。関連しない話題は、アンケートにより獲得した項目の中で、共起・共助関係に当てはまらないものの中から無作為に選択した。三つ目の話題について話している時に、三秒以上の会話の途切れがあった場合に実験終了とした。

4.2 解析

今回得られたデータから共起・共助・その他のそれぞれから一組ずつ、計三組を選び、心拍解析と音声解析の両方を試みた。

4.2.1 心拍解析

人と人がコミュニケーションする際に、生体リズムが相互に同期化する、エントレイメントという現象がある。これは元々は母子相互の働きから発見されたものだが、成人間 [渡辺 85] や人とロボット間 [飯尾 10] でも確認されている。そこで、本研究では、心拍の同期から対話における生体リズムの解析を試みる。

本実験の解析には、心拍の R-R 間隔を用いた。R-R 間隔とは、R 波の発生時刻と、一つ前の R 波の発生時刻の差を指す。R-R 間隔は、運動した時や緊張した時に間隔が狭くなり、安静状態の時は間隔が広くなる性質を持つ。心拍に基づく同期現象の研究では、R-R 間隔が指標として多く用いられている。解析方法は移動相関法を用いた。まず話題が始まってから 4 秒間の心拍間隔の相関を取る、その後開始位置を 250ms 移動させて次の四秒間の心拍間隔を取る。この動作を話題ごと会話時間全体に行い、それぞれで相関を求める。計算はピアソン相関を用いた。計算結果から、絶対値が 0.6 以上となっている 4 秒間の結果を、相関していると定義した。結果を表 1 に示す。ここで、同期回数とは四秒間以上の心拍間隔における相関に対して一回の同期と定義した。会話時間が話題毎に異なっているため、一分間あたりの平均同期回数を求めた。

実験の結果から、共起共助の関係のペアの方が、その他のペアよりも同期回数が多かった。共起ペアの場合、共通の話題である「人に教える」と「日記・ブログ」では一分間あたり 11 回以上の同期が見られるが、その他の話題では 8 回に減少している。逆に共助のペアでは最後のその他の話題が一番同期回数が多いという結果になっている。しかしながら、その他の話題は二分程度しか話しておらず、互いに興味のない話題だったことも伺える。その他のペアではフリートークの時よりも同期回数が多く、話題提示後の同期回数 10 回程度に減っている。お互いに興味がない話題であり、それでも会話を続けようと努力しているのではないかと考えられ、結果的に同期に至らなかったのではないかと推測している。しかしながら、今回の結果だけでは被験者数が少ないため、被験者の特性、例えば「人と話すのが好き」「人見知り」「口下手」などによる影響を除くことができない。今後、被験者数と組み合わせパターンを増やして解析を進める必要がある。今回の実験において、被験者へのインタビューから、話題によって話しやすいものと話しにくいものがあることがわかってきた。例えば、日記・ブログや資格などは比較的明示的であり話を広げやすいが、「人に教える」などは曖昧であり何について話せば良いのか戸惑ったという声があった。今後、どのような話題をどのように提示すれば良いのかを検討する必要がある。また今回は三秒以上の沈黙を話題の切り替わりとしたが、その閾値についても検討する必要があると感じている。そこで、次節で述べるように、音声情報からの盛り上がり盛り下がりの検出も必要であると考えいる。

4.2.2 音声信号解析

前節では心拍データを用いたエントレイメントに着目したが、ここでは、音声信号から検出できないか試みる。まず、対話中の音声データに対し、被験者の発話区間を手動で切り出した。その音声データに対して、音声解析ソフト Praat*2 を用いて、音圧実効値・最小音圧・最大音圧・音圧レンジ・平均ピッチ・最小ピッチ・最大ピッチ・ピッチレンジの 8 つの音響特徴

量を得た。

この音響特徴量を用いて、Weka [Frank 16] のニューラルネットワーク（多層パーセプトロン）を用いて分類を試みた。検証は 10-fold 交差検定で行った。分類は被験者毎に行った。ニューラルネットワークを用いたのは過去の研究 [生田 15] で同じ音声特徴量を用いた分類問題の際に精度が良かったためである。

まず表 1 の心拍解析と同様に話題に着目した。各被験者のデータに対し、話題毎に分類可能かどうか検証を行った。その結果、表 2 のように全体として 50% 以下の結果となり、両者の共通の話題について話しているのか、その他の話題なのか、推定するのは難しいことがわかった。ただし今回はフリートークも含め、さらに二つの共通話題についても別のタグを付けて分類を行っているため、完全な二値分類ではない。共通の話題か否かという二値に落とし込み再度分類を試みたい。

先の結果を踏まえ、次に会話の始まりや会話の盛り上がり、盛り下がりが検出できないか、検証を行った。三秒以上沈黙が続いた場合に新しい話題が提示されるため、話題が切り替わる箇所において、なんらかの違いが見られるのではないかと推測したからである。先ほどの音声データに、筆者らが対話内容を聞いて手作業で、話題毎の始まり、盛り上がり、盛り下がりのタグをつけた。先ほどと同様に多層パーセプトロンを用いて、被験者毎に 10-fold 交差検定で検証を行った。その結果、表 2 に示すように、60% 程度と話題毎の分類よりは高い結果を得ることができた。

またすべての被験者データを用いて分類を行った場合、73.6% という結果となった。この数値だけを見るとある程度分類ができていようにも見える。しかしながら、表 3 の混同行列をみてわかるとおり、盛り上がりのデータ数の多さに引っ張られていることが分かる。また当初目標としていた、盛り下がりの検出が全くできていないことがわかる。これは最初のデータ前処理の部分において、発話区間だけを切り取ったため、沈黙の区間が落ちているためだと考えられる。現在、全ての音響データに対して、特徴量抽出を行うことで精度をあげることができないか検証を行っている。

5. おわりに

本稿では人と人をつなぐコミュニティセントリックシステムの一つとして、CcAI について述べ、それに向けた第一歩として、人と人の会話の解析を行った。エントレイメントに着目し心拍同期について解析し、関係性の指標に使えるかどうかという見通しを得た。さらに音声特徴量を利用して、会話の始まり、盛り上がり、盛り下がりの検出を試みた。今回の前処理方法では十分な精度を得るのは難しかったが、前処理方法を再検討することで、会話の途切れなどを検出可能であると考えている。今回はオフラインによる処理だったが、ある程度の精度が得られる手法を見出し、オンラインでの関係性の観測や盛り下がりの検出を行い、ロボットがよいタイミングで話題を提供するようシステム化を進めたい。またビッグデータとの連携や人工知能のよりよい活用など様々な分野と連携を強化し、実際のサイトで運用可能なシステムとして実装を目指す。

*2 Praat : <http://www.fon.hum.uva.nl/praat/>

表 1: 心拍同期の解析結果

	話題	会話時間 (分:秒)	同期回数 (回)	1 分間当たりの平均同期回数 (回)
共起: User 4 - User 6	フリートーク	9:23	110	11.7
	共起: 人に教える	3:34	41	11.5
	共起: 日記・ブログ	7:05	81	11.4
	その他: ペットの世話	8:55	72	8.08
共助: User 2 - User 6	フリートーク	3:42	43	11.6
	共助: ゲーム	3:59	51	12.8
	共助: 人に教える	10:05	122	12.1
	その他: 掃除	2:17	35	15.3
その他: User 4 - User 10	フリートーク	8:49	112	12.7
	資格	11:09	113	10.1
	洗濯	5:59	65	10.9
	プログラミング	4:14	43	10.2

表 2: ニューラルネットワークによる分類

	被験者	話題分類正解率 (%)	話始め・盛り上がり・盛り下がり分類正解率 (%)
共起ペア	User 4	52.3	68.6
	User 6	28.4	84.1
共助ペア	User 2	52.3	66.7
	User 6	39.5	47.6
その他	User 4	37.3	55.2
	User 10	32.1	64.1

表 3: 全被験者データを用いた分類における混同行列

	始まり	盛り上がり	盛り下がり
始まり	2	112	0
盛り上がり	3	508	0
盛り下がり	0	68	0

参考文献

- [Frank 16] Frank, E., Hall, M. A., and Witten, I. H.: *The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques*, 4 edition (2016)
- [Gomi 15] Gomi, R., Suzuki, A., Sato-Shimokawara, E., and Yamaguchi, T.: Analysis of dialogue for acquiring personal characteristics toward co-occurrence matching, pp. 206–212 (2015)
- [Lee 02] Lee, W.-P., Liu, C.-H., and Lu, C.-C.: Intelligent agent-based systems for personalized recommendations in Internet commerce, *Expert Systems with Applications*, Vol. 22, No. 4, pp. 275–284 (2002)
- [Sato-Shimokawara 15] Sato-Shimokawara, E., Suzuki, A., Gomi, R., and Yamaguchi, T.: Obtaining user's preference and ability from human-robot conversation towards mutual assistance, pp. 3557–3560 (2015)
- [Sato-Shimokawara 16] Sato-Shimokawara, E., Gomi, R., and Yamaguchi, T.: Simulation and analysis of matching method for co-occurrence and mutual assistance based on personal characteristics (2016)
- [Suzuki 16] Suzuki, A., Sato-Shimokawara, E., Yamaguchi, T., and Ikehata, K.: Acquiring personal attributes using communication robots for recommendation system, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 9835, pp. 237–246 (2016)
- [Zheng 09] Zheng, Y., Zhang, L., Xie, X., and Ma, W.-Y.: Mining interesting locations and travel sequences from GPS trajectories, pp. 791–800 (2009)
- [神田 08] 神田 崇行, 塩見 昌裕, 野村 竜也, 石黒 浩, 萩田 紀博: RFID タグを用いた科学館来館者の移動軌跡の分析, *情報処理学会論文誌*, Vol. 49, No. 5, pp. 1727–1742 (2008)
- [生田 15] 生田 千紗, 下川原 英理, 山口 亨: 対話ロボットのための話題獲得に向けたキーワード検出に関する音声データ解析, pp. 1235–1238 (2015)
- [相澤 16] 相澤 秀和, 岩崎 真也, 五味 怜央奈, 下川原 英理, 山口 亨: 人と人とを結びつける対話ロボット開発に向けた人同士の会話の解析, pp. 1422–1426 (2016)
- [渡辺 85] 渡辺 富夫: 成人間コミュニケーションにおけるエンタテインメント (音声-体動同期現象) の分析, *情報処理学会論文誌*, Vol. 26, No. 2, pp. 272–277 (1985)
- [飯尾 10] 飯尾 尊優, 塩見 昌裕, 篠沢 一彦, 宮下 敬宏, 秋本 高明, 下原 勝憲, 萩田 紀博: 語彙の引き込み: ロボットは人間の語彙を引き込めるか?, *情報処理学会論文誌*, Vol. 51, No. 2, pp. 277–289 (2010)
- [本橋 16] 本橋 幸治, 下川原 英理, 鈴木 葵, 陳 苑茵, 山口 亨: 興味関心の抽出に向けた iBeacon を用いた対象地点情報および滞在時間の取得と解析, pp. 1432–1437 (2016)