

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 15 日現在

機関番号：32689

研究種目：若手研究(B)

研究期間：2013～2014

課題番号：25870824

研究課題名(和文) 多人数会話活性化ロボットのための活性化戦略

研究課題名(英文) Facilitation Strategies for Multiparty Conversation Robots

研究代表者

松山 洋一 (Matsuyama, Yoichi)

早稲田大学・付置研究所・次席研究員

研究者番号：90584467

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：多人数が参加する会話に自然に参加し、その場を活性化させ自己目的的に会話を楽しまれている状況を創る手助けをするロボットシステムの開発を行った。多人数会話においては進行中の会話についていけなかったり発話を躊躇しているような「置いてけぼり」となっている会話参加者が存在することがある。彼らは自ら発話する意欲を持ちながらも何らかの要因によって発話できないでいたり、そもそも進行中の会話に興味を失っている場合などが考えられる。本研究課題では、各参加者・参加者間の状態に応じて発話行動を変え「置いてけぼり」となっている参加者を救いながら最大多数の最大参加を促す活性化戦略やスキルを有するロボットの開発を行った。

研究成果の概要(英文)：We present a framework for facilitation robots that regulate imbalanced engagement density in a multiparty conversation with proper procedures for obtaining initiatives. In multiparty conversations, social imbalance, in which a participant is left behind in the current conversation, sometimes occurs. In such scenarios, a conversational robot has the potential to objectively observe and control situations. Consequently, we present model procedures for obtaining conversational initiatives in incremental steps to harmonize such multiparty conversations. During the procedures, a facilitator must be aware of both the presence of dominant participants leading the current conversation and the status of any participant that is left behind. We model and optimize these situations and procedures as a partially observable Markov decision process (POMDP). The results of experiments conducted to evaluate the proposed procedures show evidence of their acceptability and feeling of groupness.

研究分野：知能ロボット

キーワード：音声対話システム ヒューマン・ロボット・インタラクション

1. 研究開始当初の背景

会話ロボットは人同士の会話の媒介者として、会話をより豊かにできる可能性を持っている。例えば、通所介護施設に代表されるような高齢者ケアの現場は、様々なサービスにより来所者自身のケアが行われるだけでなく、ケアスタッフと来所者間および来所者同士の重要なコミュニケーションの場でもある。さらに、昨今ではコミュニティケアという考えに基づき、ケア施設を当該地域内の世代を超えたコミュニケーションの拠点とするような試みが行われている。豊かなコミュニケーションこそは、ケアの本質である。例えばそのような場に会話ロボットがスタッフの一員として参加し、人同士のコミュニケーションを支援したり活性化することができれば、付加価値の高いシステム/サービスになると考えられる。

我々の日常の会話は、1対1の対話だけでなく多人数で行われることが多い。音声対話システムの研究分野においては長らく1対1の対話を対象としてきたが、近年、多人数会話に拡張した研究も見られるようになってきた。申請者は、2008年より多人数会話に参加するロボットの研究を行なっており、主に高齢者ケアを目的とした多人数会話活性化システムの検討を続けてきた。実際の高齢者施設でも複数回に渡ってフィールド実験を行なっており、一定の活性化効果が認められた。しかしながら、これらの研究は、まだ永続的に会話を活性化できるシステムであるとはいえない。ロボットを介して活性化された会話の心理的側面について申請者が過去に分析した結果によれば、「楽しさ」、「静かさ」、「快適さ」などの因子が作用していることがわかった。客観的な指標としては、活性化された会話では発話量、笑顔頻度、頷き頻度などが上昇傾向を示すことがわかっている。すなわち、最大多数の参加者に対してこれらの指標ができるだけ長時間に渡って大きな値を示している状況を創り出すことが、本研究課題で実現すべき会話活性化戦略である。

2. 研究の目的

4者によって構成される多人数会話をファシリテーションするロボットのための行動戦略を提案する。2者間で行われる会話の状況では、2者間で発話権の所在や文脈が互いに了解されながら会話が進行する。従来の多くの対話システムは、このような1対1の質問応答や発話権交代モデルを想定してきた。しかし、3者以上で行われる「多人数会話」では、発話権や文脈は参加者間で一意には了解されないことが多い。これは、多人数会話における発話権あるいは参与役割(話し手、受け手、傍参与者、他)は、各参加者ごとにズレを生じながら認識される重層的なものであるからである。このような多人数会話においては、しばしば「社会的不均衡」の問題

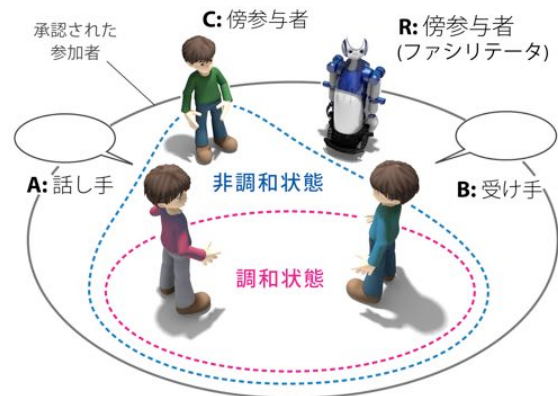


図1：4者会話の様子

を生じる。例えばあるグループを形成している3者会話において、2者間のインタラクションが主導的でもう1人の参加者が会話にうまく参加できない「置いてけぼり」(意に反して、参与役割における「話し手」にも「受け手」にも長時間昇格できない)状態になっているような状況を考えてみる。このとき、「置いてけぼり」状態の参加者は本来自分の考えを発言したいにも関わらず何らかの原因によってその機会をうまく与えられていないような状況に陥っている可能性がある。さらに、このような不均衡の問題は、多人数会話の文脈の重層性という特性上、会話を主導的に進めている参加者側からは了解されていない場合もあり、その場合は自然には解決しない可能性が高い。

このような社会的な不均衡を解決するためには、その状況を調整する役割を担う第4番目の参加者を考える必要がある。物理的に場を共有する会話ロボットは、その第4番目の参加者の役割を担える可能性を有している。図1は、ロボットを調整役として参加させた4者会話の様子である。ここで注意すべきは、ロボットがいざ場の調整に乗り出そうとするときにはそれ相応の手続きが必要になることである。解決すべき「置いてけぼり」状態が検出された時点で即刻直接行動に出してしまうと、会話の場を破壊してしまう可能性があるからである。会話の場を乱すことなく調整をするためには、その場で主導的に会話を進めている参加者らの状況も考慮しつつ、場を調整するための主導権を要求すべく行為し、明示的あるいは暗黙的にそれが承認された後に、改めてしかるべき対象者に発話機会を与える、といったような多段の手続きを踏む必要がある。

本論文では、ファシリテーションモデルを扱うための最小単位である(ロボットを含む)4者会話を対象として、ロボットの会話状況の調整手続きについて検討する。この手続きのモデル化には、誤りを含むセンサ情報にロバストな部分観測マルコフ決定過程(POMDP)を用いる。

3. 研究の方法

3.1 主導権奪取手続き

ある時点において主導的に会話を進めている参加者から主導権を委譲されるためには、一度、主導者が展開している会話に参加して、その内部から次の展開を開始する必要がある。さもなくば、会話の場が分裂してしまう恐れがあるからである。例えば図1のような状況で、AとBが会話している様子を無視してロボットがCに話しかけてしまうと、A-Bの会話とロボット-Cの会話は分裂したものになってしまうだろう。そうならないためには、ロボットが一度AとBの間で行われている会話に一旦参加し、次の主体的な会話の展開を暗黙的に許されるような状況をつくる必要がある。ここではこのようにある時点の主導的な会話に参加している状態を「調和状態」と呼ぶことにする。すなわち、参与役割における「話し手」と「受け手」は自動的に調和状態にある。しかし、「傍参与者」は状況によって調和状態と非調和状態に分かれる。「傍参与者」が調和状態になるためには、上の例におけるAとBが生成する隣接ペアに呼応する必要があると考えることにする。以上を踏まえて、本論文では主導権奪取手続きを、調和状態の参加者（以下、調和者）と非調和状態の参加者（非調和者）ごとの行動の制約として次のように定義する。

(1) 発話権譲渡先に関する制約：非調和者は、他の非調和者に対して直接話を振ってはいけない。

(2) 話題変更に関する制約：調和者は、非調和者に対して話題変更を行なってはいけない。

本研究では、これらの制約を、強化学習(POMDP)の枠組みの中でモデル化する。

3.2 「置いてけぼり」状態の推定

参与役割の視点から、「置いてけぼり」の状態について考える。図3のような状況で、AとBが活発に会話を行い、Cに発話機会がない状態であったとする。このとき、Cは傍参与者である時間帯が結果的にしばらく続いていることになる。また、このような「置いてけぼり」状態は、話題が変わると改善される場合がある。本論文では、各参加者が当該話題区間における傍参与者にアサインされている割合を、「置いてけぼり度(=傍参与状態の深さ $Depth_{SPT}$)」と定義する。

$$Depth_{SPT} = (Duration_{SPT}) / (Duration_{topic})$$

3.3 会話参加意欲度

「置いてけぼり」状態にある参加者は、さらに会話参加への意欲をもっているか否かで状態が分かれると考えられる。すなわち、発話機会が与えられていないが自分も

積極的に参加しようと思っている参加者と、そもそも現在の会話にあまり興味がない場合である。我々がロボットを介して活性化された会話(各参加者の参加意欲が高い状態にある会話)の心理的側面について分析した結果によれば、「楽しさ」、「賑やかさ」、「快適さ」などの因子が作用していることがわかっている。客観的な指標として、そのような心理状態にあるときに各参加者の発話量、笑顔頻度、頷き頻度などが上昇傾向を示すこともわかっている。よって本研究課題では、参加者*i*の活性化度を、発話量 $SpeechAmount_i$ 、笑顔量 $SmilingAmount_i$ 、頷き時間 $NoddingAmount_i$ から以下のように計算されるものと定義する。

$$Activity_i = \alpha(SpeechAmount)_i + \beta(SmilingAmount)_i + \gamma(NoddingAmount)_i$$

3.4 POMDP による手続きの最適化

会話活性化戦略の効果を最大化させるために、部分観測マルコフ決定過程 (Partially Observable Markov Decision Process: POMDP) を用いてモデル化する。

一般に、POMDP は式 (3) のように状態 S 、システム行動 A 、状態遷移確率 T 、報酬 R 、観測確率 Z 、割引率 γ 、初期の信念状態 b_0 の集合として定義される。

$$\beta = (S, A, T, R, Z, \gamma, b_0)$$

信念状態は以下のように更新される。

$$b'(s') = \gamma \cdot P(o'|s', a) \sum_s P(s'|s, a) b(s)$$

ここで、 S を各参加者の調和状態を S_h 、「置いてけぼり」状態の参加者の参加意欲を S_m 、参加者の行動を A_p を用いて定義する。

$$s = (s_h, s_m, a_p)$$

すると、状態の遷移確率は以下のように表現できる。ここで、簡単のため状態遷移確率に

$$\begin{aligned} P(s'|s, a) &= P(s'_h, s'_m, a'_p | s_h, s_m, a_p, a_s) \\ &= P(s'_h | s_h, s_m, a_p, a_s) \cdot \\ &\quad P(s'_m | s'_h, s_h, s_m, a_p, a_s) \cdot \\ &\quad P(a'_p | s'_m, s'_h, s_h, s_m, a_p, a_s) \end{aligned}$$

次のような仮定を置く。

$$T_{S_h} = P(s'_h | s_h, a_p, a_s) \quad (\text{調和状態モデル})$$

$$T_{S_m} = P(s'_m | a_s) \quad (\text{参加意欲モデル})$$

$$T_{A_p} = P(a'_p | s'_h, a_p, a_s) \quad (\text{参加者行動モデル})$$

よって、信念状態の更新式は以下のようになる。

$$b'(s'_a, a'_p) = \gamma \cdot P(o' | s'_h, s'_m, a'_p, a_s) \cdot \sum_{s_h} P(s'_h | s_h, a_p, a_s) \sum_{s_m} P(s'_m | a_s) \cdot \sum_{a_p} P(a'_p | s'_h, a_p, a_s) \sum_{a_p} b(s_a, a_p)$$

調和状態，ユーザ行動，システム行動はそれぞれ表 1~3 のようになる。状態遷移図は図 4 のようになる。

表 1：調和状態

調和状態とその意味	
Unharmonized	ロボットは現在の会話に調和していない
Pre-harmonized	ロボットは現在の会話に調和する為の承認を待っている
Harmonized	ロボットは現在の会話に調和している

表 2：参加者行動

参加者行動とその意味	
first	参加者が第 1 部分を発した (Question)
second	参加者が第 2 部分を発した (Answer)
third	参加者が第 3 部分を発した
other	参加者が、ロボット以外の人に発話を行なった
call	参加者がロボットを呼んだ

表 3：システム行動

ロボットの取りうる行動とその意味	
answer	現在の話し手の質問に対して答える
question-new-topic	新たな話題で質問する
question-current-topic	現在の話題で質問する (話をふる)
trivia	トリビアを言う
simple-reaction	呼ばれたことに関して反応する
nod	現在の話者に頷く
none	何もしない

また、報酬として、3.1 で考察したファシリテータとしての行動の制約も加味し、行動の適切性の観点から設定した。

3. 5 実験

本論文で提案した主導権奪取手続きを評価するために、手続きの適切性と場の一体感の評価 (実験 1)、手続き開始タイミングの適切性の評価 (実験 2) について被験者実験を行った。音声認識誤りなどの予めロボットを含む 4 者会話 (A, B, C, Robot) を収録したビデオを視聴してもらふことにより評価実験を行った。

実験 1：主導権奪取手続きの適切性と一体感

- 条件 1：手続き無し (話題変更無し)
A と B のひと続きのインタラクションが隣接ペア第 3 パートによって区切られた直後に C に直接質問するといったように、主導権奪取手続きを用いることなく、直接的に発話権の制御を行った場合。このとき、既出の話題と同じもの (「もののけ姫」) について C に質問している。

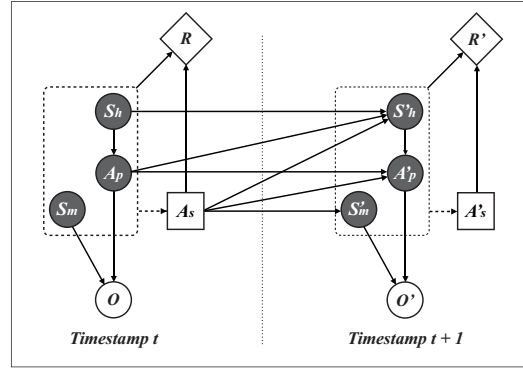


図 2：状態遷移図

- 条件 2：手続き有り (話題変更無し)
A と B のひと続きのインタラクションが隣接ペア第 3 パートによって区切られた直後に、一旦 A に第 1 パートで話しかけ、A からの反応 (第 2 パート) を待ち、A とのインタラクションを区切ってから C に改めて発話の機会を与える。このとき、条件 1 と同様に既出の話題と同じもの (「もののけ姫」) について C に質問している。

- 条件 3：手続き無し (話題変更有り)
条件 1 におけるロボットの C への質問において、新規話題を用いるもの。

- 条件 4：手続き有り (話題変更有り)
条件 2 におけるロボットの C への質問において、新規話題を用いるもの。

各ムービーは、それぞれ 30 秒程度である。各条件のムービーを視聴した後、被験者は (a) ロボットの参加者としての行動の適切性と (b) 一連の会話によって生じる一体感について、7 段階のリッカートスケールのアンケートを求められた。

実験 2：手続き開始タイミングの適切性

比較する条件として、第 1 部分、第 2 部分、隣接ペアの途中で手続きを開始するものを用意した。ここで、第 3 部分のタイミングで手続きを開始することは適切であると予想されるため、今回は比較していない。

- 条件 1：第 1 部分の直後で手続き開始
- 条件 2：第 2 部分の直後で手続き開始
- 条件 3：隣接ペアを考慮せずに手続きを開始

3 条件のムービーを視聴した後、被験者はロボットの参加者としての行動の適切性について、7 段階のリッカートスケールのアンケートを求められた。

結果

実験 1 の結果を図 3 と 4、実験 2 の結果を図 5 に示す。

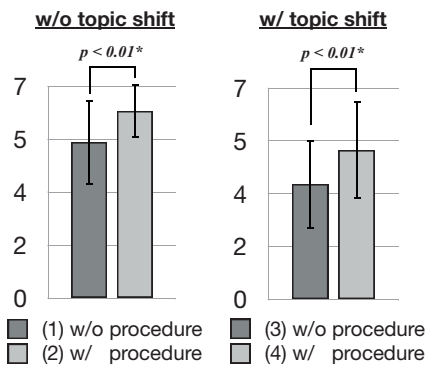


図3：実験1（適切性）の結果

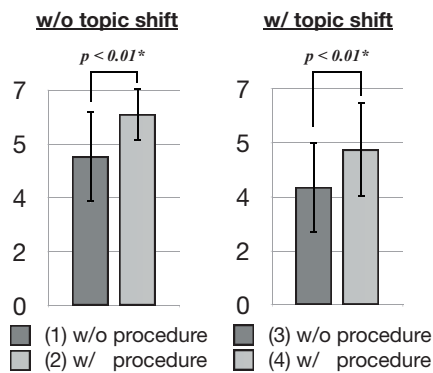


図4：実験1（一体感）の結果

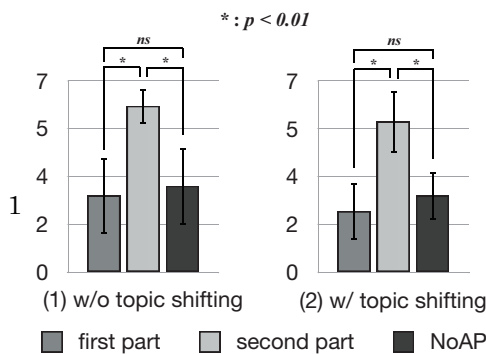


図5：実験2（手続き開始のタイミング）の結果

図3より、話題変更を行わない場合も、行う場合も、手続きをとったほうがより適切であるということがわかる。t検定を実施したところ2条件間に優位な差が見られた ($p < 0.01$)。また、図4より、手続きをとった時のほうが、手続きを取らない時よりも、一体感を生むことがわかった。t検定を実施したところ2条件間に優位な差が見られた ($p < 0.01$)。

また、手続きを開始するタイミングについて、図5を見ると、トピックを変更する、しないに関わらず、第2部分で手続きを開始することが適切であることが示されている。話題変更を伴わない手続きを行うタイミングでは、分散分析の結果、行動選択の適切性について条件間で有意な差が認められた ($F[2,26]=34.456$, $p < 0.01$)。

これらの結果から、主導権奪取手続きを用いて「置いてけぼり」状態の参加者にアプローチする方が、一会話参加者としてより適切な行為であり、それによって場の一体感が生まれやすくなることが示唆されたと言える。また、その手続きを開始するタイミングについては、隣接ペアの第1部分直後よりも第2部分直後に行う方が会話参加者にとってより適切であると感じられることがわかった。

4. 研究成果

本研究課題では、多人数会話において発生する発話機会の不均衡を調整するために、機会が十分に与えられず「置いてけぼり」状態になっている会話参加者に、会話ロボットが適切な手続きを経て発話機会を渡すための計算モデルを提案した。

特に、ファシリテーションのモデルを扱うための最小単位である（ロボットを含む）4者会話を対象している点が本研究の特徴である。会話の場で主導的に会話を進めている参加者らの状況も考慮しながら段階的に主導権を奪取ししかるべき対象者に発話機会を与えるような手続きの提案は、全く新しいものである。

本成果は、音声言語処理分野で国際的にインパクトの大きいジャーナル論文誌 IEEE/ACM Transactions on Audio, Speech, and Language Processing, および Journal of Computer Speech and Language に採録されている。

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計2件）

(1) Yoichi Matsuyama, Akihiro Saito, Shinya Fujie and Tetsunori Kobayashi, Automatic Expressive Opinion Sentence Generation for Enjoyable Conversational Systems, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2014. (DOI:10.1109/TASLP.2014.2363589)

(2) Yoichi Matsuyama, Iwao Akiba, Shinya Fujie and Tetsunori Kobayashi, Four-Participant Group Conversation: A Facilitation Robot Controlling Engagement Density As the Fourth Participant, Journal of Computer Speech and Language, 2014. (DOI:10.1016/j.csl.2014.12.001)

〔学会発表〕（計5件）

(1) Yoichi Matsuyama, Tetsunori Kobayashi, Towards a Computational Model of Small Group Facilitation, AAAI 2015 Spring Symposia Turn-taking and Coordination in

Human-Machine Interaction, March 2015.
(査読有り)

(2) Yoichi Matsuyama, Iwao Akiba, Akihiro Saito and Tetsunori Kobayashi, A Four-Participant Group Facilitation Framework for Conversational Robots, Association for Computational Linguistics, Proceedings of the SIGDIAL 2013 Conference, Metz, France, pp.284-293, August 2013.
(査読有り)

(3) 松山洋一, 中川純, 渡井大巳, 林明宏, 遠田敦, 和田康孝, 行動をデザインする: 人の行動を促す人間-環境インタラクシオンデザイン, 情報処理学会論文誌 サイバーフィジカルシステム特集号, 2014. (招待論文)

(4) 松山洋一, Alexandros Papangelis, Ran Zhao and Justine Cassell, 2 者会話におけるラポール形成・維持・崩壊の計算モデル, 人工知能学会 言語・音声理解と対話処理研究会, SIG-SLUD, December 2014. (査読無し)

(5) 松山洋一, 齋藤彰弘, 小林哲則, 意外性のある質問応答システムのための意見文自動生成と発話文組み合わせ手法, 日本音響学会 2013 秋季研究発表会, NO.3-8-2, September 2013. (査読無し)

[その他]

ホームページ等

www.pcl.cs.waseda.ac.jp

6. 研究組織

(1) 研究代表者: 松山 洋一

(MATSUYAMA YOICHI)

早稲田大学 GCS 研究機構・次席研究員

研究者番号: 90584467