

多人数会話ファシリテーションロボットの 主導権奪取手続き

秋葉 巖^{1,a)} 松山 洋一¹ 小林 哲則¹

概要: 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適切な手続きを経て発話機会を与えることのできる会話ロボットを提案する。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4者会話(ロボットを含む)を対象として、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計算モデルを提案する。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。さらに手続きとそのタイミングの適切性について評価した被験者実験の結果を報告する。

キーワード: 多人数会話, 音声対話システム, ファシリテーション, 主導権, 隣接ペア

Abstract: We propose a facilitation robot harmonizing four-participant conversations. Four-participant conversation is the minimum unit that needs facilitation skills. In general, three is the minimum number of participants of a multiparty conversation. In such three-participant situations, back-and-forth interactions between two participants out of three primarily occur and another participant tends to be left behind, who cannot properly get floors to speak. Here, they need one more participant who helps the participant left behind to harmonize him/her with the others. Conversational robots have potentials to participate in such conversations as the fourth participant. When the robot steps in the situation to help, there should be proper facilitating procedures to obtain initiatives to control conversational contexts. In this paper, we propose a conversational robot system harmonizing four-participant conversations along procedures of obtaining initiatives of topic and floor control. These situations and procedures were modeled and optimized as the partially observable Markov decision process. We conducted an experiment to evaluate appropriateness of the proposal procedures and the result shows evidence of its acceptability.

Keywords: multiparty conversation, spoken dialogue systems, facilitation, initiative, adjacency pair

1. はじめに

本論文では、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。図1(a)は2者間で行われる会話の様子を示している。このような状況では、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[15]。しかし、図1(b)のような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[7][5]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「社会的な不均衡」の問題を生じる。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

¹ 早稲田大学 情報理工学研究所
Major in Computer Science, Graduate School of
Fundamental Science and Engineering, Waseda University
27 Waseda Shinjuku-ku, Tokyo, Japan
^{a)} akiba@pcl.cs.waseda.ac.jp

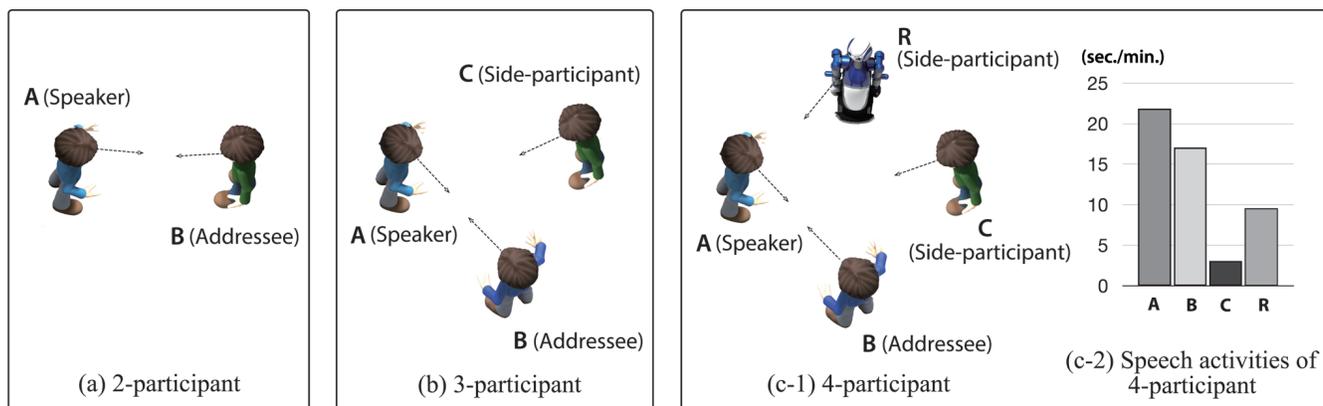


図1 (a) 2者間会話モデル, (b) 3者間会話モデル (多人数会話の最小単位), (c-1) 4者間会話モデル (ファシリテーションモデルの最小単位). (c-2) ある区間における4者の発話量の例.

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(c-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。図1(c-2)は、ある会話の区間における各参加者の発話量の例である。この例では、参加者Cの発話量が相対的に少なく、しばらく「置いてけぼり」状態である可能性が高いと考えられる。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。

多人数会話に参加するエージェントを取り扱った先行研究として、松坂らは物理的にロボットを多人数会話に参加させた先駆的な研究を行った[10]。松山らは、高齢者施設で日常的に行われているクイズゲームに参加者の一人としてロボットを参加させ、ヒントを提供したりクイズの解答に応じた蘊蓄などを披露することで会話の場をより楽しいものにさせるシステムを開発し[12]、そのようなロボットの存在による効果を実験を通して確認した[11]。堂坂らは、クイズ会話をタスクとして思考喚起型の多人数会話システムを開発した。エージェントの存在と共感的な発話が、参加者のクイズに参加したことへの満足度を向上させ、結果的に発話量も増加したと報告されている[6]。Bohus et al.は、不特定多数のユーザが訪れるキャラクタエージェントを用いて、ユーザの視線やジェスチャー、音声などのマルチモーダルな入力情報を用いたファシリテーションモデルの先駆的な検討を行った[4]。一方ファシリテーションの方法については、1940~1950年代ごろから社会心理学の分

野で Benne et al. [3] や Bales et al. [2] が小規模なグループ内のダイナミクス (Small Groups Dynamics) について先駆的な検討を行った (以下, Small Groups を単に「グループ」と訳す)。彼らは、グループでのファシリテーターの行為の機能的側面 (Functional Roles of Facilitation) と感情的側面 (Socio-Emotional Roles) について分析した。Kumar et al. は、この Bales らの整理したファシリテーターの社会感情的側面 (Socio-Emotional Interaction Categories) に注目して、テキストベースのチャットシステムを用いて、ユーザに共感したりするような発話行為の効果についてユーザ評価実験を行った[8]。しかしながら、ファシリテーションのためのエージェントの行動戦略の計算モデル、特に上に述べたような発話機会の不均衡を調整するための適切な手続きのモデルについてはほとんど検討されてこなかった。

よって本論文では、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。以下の章では、まずグループの調整についての先行研究を整理し、本研究の立場を明らかにする。3章では、会話状況とそれに基づいた調整手続きを POMDP を用いてモデル化し、4章でシステム・アーキテクチャの全体について紹介する。5章で提案する手続きの適切性についての被験者実験の結果について述べ、6章で結論と今後の展望について述べる。

2. グループの調整

Benne et al. はグループ内での個人の行動をより理解するために、各グループ構成員の機能的役割 (Functional Roles) について分析した[3]。機能的役割はタスク遂行役割 (Task Role)、グループ調整役割 (Group Maintenance Role)、個人的役割 (Individual Role) に分類される。タスク遂行役割は、グループで取り組まれるタスクに関連する役割と定義

され、タスクを完了するために必要な協力的な行動などを対象とした。グループ調整役割は、グループがグループとして機能するための役割と定義され、グループ内の個人個人の関係性に関する行動を対象とした。個人的役割は、参加者自身が満たしたい目的に関連する役割と定義され、グループの目的やグループ内の個人個人の関係性に関するものない行動を対象とした。

Bales et al. は、Benne et al. の研究を引き継ぎ、1950 年にグループ内での行動をタスク達成側面 (Task Area) と感情的側面 (Socio-emotional Area) に分けた。感情的側面では、グループ内での関係に関するものが取り扱われており、例えば会話の主導権を取る、誰かの発言を補足するなどがそれに当たる。Bales et al. と Benne et al. が行ったグループ内の行動の分類は、多くの研究で利用され評価されている [13] [1]。また、Bale と Benne の分析に基づき、タスク遂行役割がグループのタスクに及ぼす影響の研究も多くなされてきた [9]。しかし、具体的にどのような手続きでグループを調整すべきかについては、ほとんど言及されてこなかった。よって本研究では、ロボットによってグループを調整するための手続きについて検討する。それには、いくつかの必須のスキルが存在する。Benne et al. の研究ではグループ調整役割は 7 つに分類されているが、本論文では概念の比較的近いものをまとめて 3 つの抽象的な役割に分類し、それらを兼ねそそえた計算モデルを記述する。

- (1) トピック調整役割 (Topic Maintenance Role): この役割は、トピックの選択に関する役割である。この役割の人は、話題の提案をしたり、話題がわからなくなっている人がいたら手助けをしたりする。Bales の役割でいうところの compromiser, harmonizer, standard setter である。
- (2) 発話権調整役割 (Floor Maintenance Role): この役割は、発話権に関する役割である。この役割の人は、発話権がしばらくの間渡っていない人に、質問をするなどにより発話権を渡し、グループ参加者全員の発話機会を均等にしようとする。Bales の役割でいうところの gatekeeper, expediter, encourager である。
- (3) 会話状況把握役割 (Observation Role): この役割は、会話の状況を把握する役割である。現在のトピックがどのようなものであるのかを把握したり、各参加者のムードや参加意欲、各参加者同士の関係性を把握しようとする。Bales の役割でいうところの Observer, commentator, Follower である。

3. グループ調整手続きの最適化計算

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

表 1 発話権を譲渡する際の制約

主体 \ 対象	調和者	非調和者
調和者	permitted	permitted
非調和者	permitted	NOT permitted

表 2 話題変更の際の制約

主体 \ 対象	調和者	非調和者
調和者	permitted	NOT permitted
非調和者	NOT permitted	NOT permitted

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 (c-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 と B の会話とロボットと 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度主体的に参加し、次の発言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 [16] に呼応する必要があると考えることにする。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。

この制約についての主体と対象の関係を、表 1、表 2 に示す。この制約を順守することによって、会話を破壊せずに発話権や話題の制御を行うことが可能になる。すなわち、この制約の順守が主導権奪取の手続きを取ることを等価である。

3.2 グループ会話状況の表現

調和 / 非調和 状態および、会話参加に対する参加意欲の記述について考える。例えば図 1 (c-1) において、C がある一定時間発話権をとることができないことが続いたとする。このような状況は、話題が変更される事により C が会話に参加しやすくなり「置いてけぼり」状態が解消される可能性があると考えられるので、同一話題の継続区間における傍参与者である割合を、傍参与状態の深さ $Depth_{SPR}$ 、すなわち「置いてけぼり」度とし、この度合いがある閾値を上回ったときに非調和状態に陥ることとする。

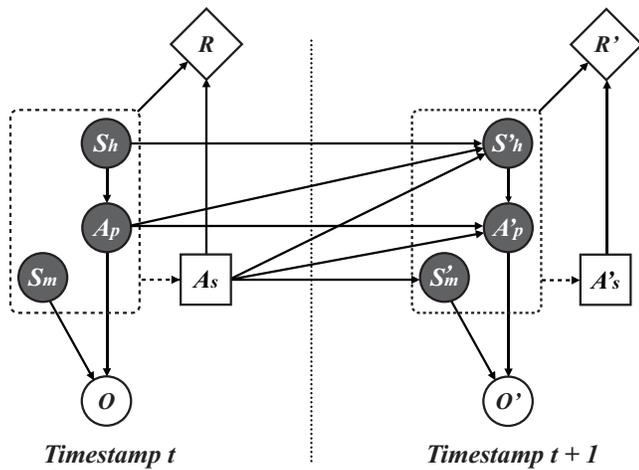


図2 POMDP の状態遷移図．各円は，確率変数を表す．グレーの円は潜在的な確率変数を表し，白抜きは観測可能な変数である．四角は，システム行動，．菱形は，状態とシステム行動の結果から与えられる報酬を表す．

$$Depth_{SPT_i} = Duration_{SPT_i} / Duration_{topic_j} \quad (1)$$

$$UnharmonizedSPT = \begin{cases} SPT_i & \text{if } Depth_{SPT_i} > Threshold \\ none & \text{otherwise} \end{cases} \quad (2)$$

なお本論文では，各発話毎に必ず 1 人の話し手と 1 人の受け手が存在し，1 人以上の傍参加者が存在すると仮定する．

3.3 POMDP による手続きの最適化

上記で表した手続きを最適化するために，部分マルコフ決定過程 (POMDP) を用いてモデル化する [18]. POMDP は， $\beta = \{S, A, T, R, O, Z, \gamma, b_0\}$ の集合として表現される．ここで， S はこの POMDP で取り扱う状態の集合， A はロボットが取りうる行動の集合， T はある状態から次の状態への遷移確率 ($P(s'|s, a)$)， R は期待報酬 $r(s, a)$ ， O は観測されるユーザー行動の集合， Z は $P(o'|s', a)$ で表される観測確率， γ は割引率 ($0 < \gamma < 1$)， b_0 は初期状態を表す．各時点において，信念状態 b は，以下のように更新される．

$$b'(s') = \gamma \cdot P(o'|s', a) \sum_s P(s'|s, a) b(s) \quad (3)$$

本研究では， S は 3 つの状態に分ける．参加調和状態 S_h ，参加意欲状態 S_m ，参加者行動 A_p で表す．よって，POMDP の状態 S は，

$$s = (s_h, s_m, a_p) \quad (4)$$

として表し，信念状態は

$$b = b(s_h, s_m, a_p) \quad (5)$$

と表す．

信念状態の更新においては，経験則に基づきいくつかの仮定を置く．

$$\begin{aligned} P(s'|s, a) &= P(s'_h, s'_m, a'_p | s_h, s_m, a_p, a_s) \\ &= P(s'_h | s_h, s_m, a_p, a_s) \cdot \\ &\quad P(s'_m | s'_h, s_h, s_m, a_p, a_s) \cdot \\ &\quad P(a'_p | s'_m, s'_h, s_h, s_m, a_p, a_s) \end{aligned} \quad (6)$$

図 2 に状態遷移図を示す．式 6 の第 1 項を調和状態モデル (*participants' harmony model*) T_{S_h} ，第 2 項を参加意欲モデル (*participants' motivation model*) T_{S_m} ，第 3 項をユーザー行動モデル (*participants' action model*) T_{A_p} と呼ぶことにする．

調和状態モデル T_{S_h} は，現会話に調和しているかどうかを表している．調和状態は，前時刻の調和状態，ユーザー行動，システムのみによって毎回決定されると仮定し，

$$T_{S_h} = P(s'_h | s_h, a_p, a_s) \quad (7)$$

と表現する．また，表 3 に示すように 4 者会話におけるロボットの主導権を奪取の手続きを考えるためにはロボットの調和状態のみ考慮すれば良いと仮定する．調和状態の遷移確率は，前述した制約に基づいて次のように遷移するよう定義した．例えば，非調和状態であるロボットが現在の話し手に質問をした場合，*Pre*-調和状態になり，調和状態になるための承認を待つ．現在の話し手がロボットに対して返答した場合，ロボットは調和状態になるとする．もし，このとき話し手がロボットに対して返答しなかった場合は，ロボットは非調和状態に戻ることとする．一度調和状態になった後で，ロボットが発話を行わない場合は徐々に非調和状態になるようにする．

参加意欲モデル T_{S_m} は，非調和者の参加意欲を表現するものである．参加意欲は，前時刻のロボットの行動によってのみ影響されるとする．参加意欲は，非調和者の ID と (true/false) の 2 値で表される．

$$T_{S_m} = P(s'_m | a_s) \quad (8)$$

ユーザー行動モデル T_{A_p} は，前時刻のユーザー行動と，以前のロボットの行動，現在のロボットの調和状態に影響を受けると仮定し，

$$T_{A_p} = P(a'_p | s'_h, a_p, a_s) \quad (9)$$

と表す．表 4 に表すように，ユーザー行動は，隣接ペアのタイプも表現している．

また，観測確率 Z を以下のように定義する．

$$Z = P(o'|s', a) = P(o'|s'_m, a'_p, a_s) \quad (10)$$

以上の定義により，信念状態の更新式は以下のように書ける．

表3 調和状態 S_h

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4 参加者行動 A_p

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5 システム行動 A_s

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

$$b'(s'_m, a'_p) = \gamma \cdot \underbrace{P(o'|s'_m, a'_p, a_s)}_{\text{observation model}} \cdot \underbrace{P(s'_m|a_s)}_{\text{motivation model}} \cdot \underbrace{P(a'_p|s'_h, a_p, a_s)}_{\text{participants' action model}} \cdot \sum_{s_h} \underbrace{P(s'_h|s_h, a_p, a_s)}_{\text{harmony model}} \cdot b(s_m, a_p) \quad (11)$$

表5に、ロボットが取りうる7種類の発話行動を示す。報酬は、前述の手続きを満すために行動の適切性の観点から与えられた。

4. システムアーキテクチャ

図3にシステム・アーキテクチャの全体図を示す。アーキテクチャは主に状況理解モジュール (Utterance & Situation Analysis), 対話制御/グループ調整モジュール (Dialogue Management), 応答文/質問文生成モジュール (Question/Answer Generation) から構成される。状況理解モジュールは、参与役割推定器、隣接ペア認識器、質問文解析器から構成される。参与役割推定器は、顔向きと発話区間の情報を用いて話し手、受け手、傍参与者の役割識別をする本論文の実験では、参与役割は手書きルールによって構成された推定器を用いている。話し手と話し手の認識精度はそれぞれ、75.1%と67.2%である。隣接ペア認識器では、音声認識から抽出されるキーワードの結果に基づいて手書きのルールによって識別する。質問文解析器では、音声認識結果に含まれる5WHの疑問詞と述部から、その質問が事実を問うFactoid型質問なのか、理由や事象の説明に基づく正答を求めるNon-factoid型質問かを識別し、応答すべき発話文タイプを決定する。対話制御/グループ調整モジュールは、3章で述べたPOMDPの方策に基づいて発話行

動を決定する。応答文/質問文生成モジュールでは、具体的な発話文を生成する。実験プラットフォームとしては、我々の開発した会話ロボットSCHEMA(シェーマ)を用いた。シェーマは、高さが、大人が座った時の目線と同じ約1.2mで、両目、眉毛、首に10自由度を持つ。また、各腕にそれぞれ6自由度の関節を持ち、ジェスチャーを表現する。口にも1自由度あり、ロボットの発話を表現する。音声認識器を含むそれぞれのモジュールは、我々が開発したミドルウェアMONEA(Message-Oriented Networked-robot Architecture)によって同期が取られている[14]。

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評価対象外の要因を排除するために、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うことにより評価実験を行った。被験者には、ビデオ視聴前に次のように説明された。「AとBは親しい関係にあり、Cは初めてこの場にやってきた人物であるため、場になかなか馴染めないという設定です。また、Cとロボットはビデオの冒頭から「置いてけぼり」状態になっています。このような状況でロボットは場を調和させるように行います。場が調和している状況とは、発話機会が会話参加者全員にできるだけ与えられている状況であるとします。」

5.1 実験1: 主導権奪取手続きの適切性と一体感

被験者は早稲田大学の学生35名(男性23名, 女性12名), 平均年齢は20.5歳だった。比較する条件として以下の4つを準備した。

- 条件1: 手続き無し(話題変更無し)

図7に示すように、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。

- 条件2: 手続き有り(話題変更無し)

図8に示すように、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(#6)において、

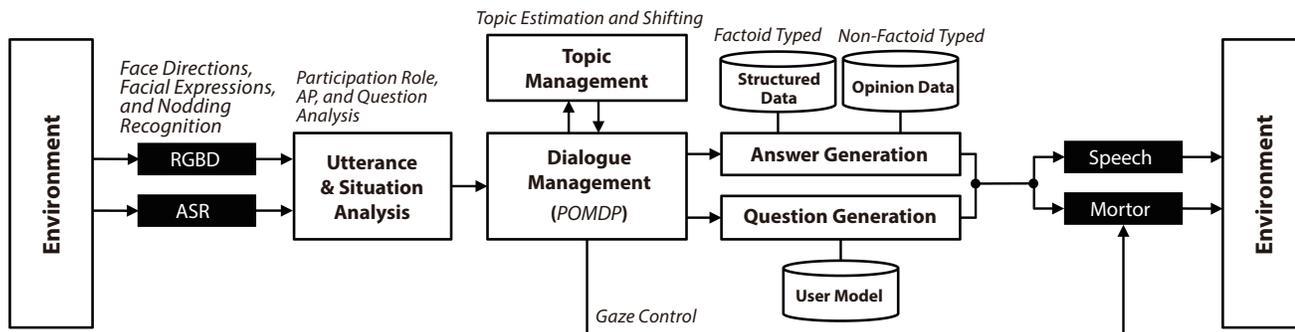


図3 システム・アーキテクチャ．状況理解モジュール (Utterance & Situation Analysis) , 対話制御/グループ調整モジュール (Dialogue Management) , 応答文/質問文生成モジュール (Question/Answer Generation) から主に構成される．

新規話題を用いるもの(ここでは「コクリコ坂から」)．

- 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(#7)において, 新規話題を用いるもの(ここでは「コクリコ坂から」)．

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．

5.2 実験2: 手続き開始タイミングの適切性

被験者は早稲田大学の学生32名(男性21名,女性11名),平均年齢は20.5歳だった．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．

5.3 結果

実験1-1の結果を図5,実験1-2の結果を図6,実験2の結果を図9に示す．図5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6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9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,(a)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($F[2,26] = 34.456, p < 0.01$)．さらにTukeyHSD法による多重比較を行ったところ,条件1と条件2,条件2と条件3で有意な差が見られた($p < 0.01$)．また,話題変更を伴う手続きを行うタイミングでも,分散分析の結果,(a)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($F[2,26] = 42.520, p < 0.01$)．さらにTukeyHSD法による多重比較を行ったところ,条件1と条件2,条件2と条件3で有意な差が見られた($p < 0.01$)．

6. まとめ

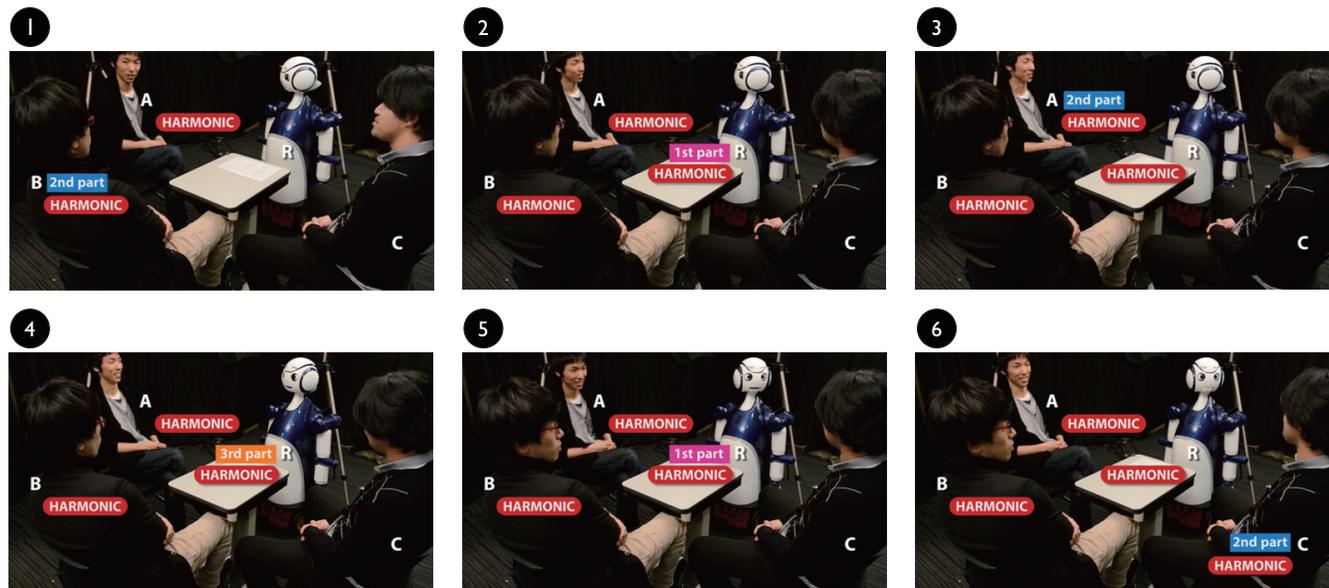
多人数会話において発生する発話機会の不均衡を調整するために,機会が十分に与えられず「置いてけぼり」状態になっている会話参加者に,会話ロボットが適切な手続きを経て発話機会を渡すための計算モデルを提案した．ファシリテーションのモデルを扱うための最小単位である(ロボットを含む)4者会話を対象して,会話の場で主導的に会話を進めている参加者らの状況も考慮しながら段階的に主導権を奪取し,しかるべき対象者に発話機会を与えるよ

#	SPK → ADD	AP	Sentences
1	A→B	1st	ものけ姫,見たことありますか?
2	B→A	2nd	あ,あります.
3	A→B	1st	そうなんですか.
4	B→A	2nd	はい.
5	A→B	3rd	へえ.
6	R→C	1st	ものけ姫は見たことある?
7	C→R	2nd	あります.

図7 条件1のトランスクリプト

#	SPK → ADD	AP	Sentences
1	A→B	1st	ものけ姫,見たことありますか?
2	B→A	2nd	あ,あります.
3	A→B	3rd	そうなんですか.
4	R→A	1st	僕も,ジブリの中でも好きな作品だよ.
5	A→R	2nd	へえ.
6	R→A	3rd	うん.
7	R→C	1st	ものけ姫は見たことある?
8	C→R	2nd	あります.

図8 条件2のトランスクリプト



#	SPK → ADD	AP	S_h	Sentences
				(Topic: "007 スカイフォール")
1	A→B	1st	Un	007の話なんですけど.
2	A→B	1st	Un	「スカイフォール」観たことありますか？
3	B→A	2nd	Un	いや、観てないですね. ①
4	A→B	3rd	Un	あ、観てないですか.
5	R→A	1st	Pre	あのね、ボンドガールの設定もとても良いよね. ②
6	A→R	2nd	Pre	そうなんだ.
7	R→A	1st	Pre	「時代遅れ」や「古い」というテーマも、44歳のジェームズ・ボンドに 対して良かったと思うんだ. ③
8	A→R	2nd	H	そうなんだ. (ロボットは主導権を持つことを暗黙的に許可されたと理解) ④
9	A→R	3rd	H	うん. ⑤
10	R→C	1st	H	「スカイフォール」は観たことある？ ⑥
11	C→R	2nd	H	観たことない.
12	A→C	1st	H	あ、観たことないですか.
13	C→A	2nd	H	全然無いんですよ.

図4 実験での会話の例。「AP」は隣接ペア (Adjacency Pair) を表す。#4 では、システムは参加者 A の隣接ペア第 3 パートを認識し、自発的な意見を第 1 パートとして A に向かって発話した (#5)。このとき、システムは自己の調和状態 S_h が *Un-Harmonized* から *Pre-Harmonized* に変わったと理解している。その後、A の (システムに対する) 第 2 パートを観測されたことをもって (#8)、調和の状況は *Harmonized* 状態となり、場を調整する主導権を奪取できたと理解している。しかるのち、ロボットは C に対して質問行動を取ることで、発話権を譲渡している (#10)。

うな手続きを提案した。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タな部分観測マルコフ決定過程 (POMDP) を用いた。さらに手続きとそのタイミングの適切性について評価した結果、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。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。

謝辞

本研究は、日本学術振興会 科学研究費 若手研究 (B)「複数人会話活性化ロボットの開発と評価」(課題番号 23700239) の補助により行われたものです。また、ロボットの開発にあたり、本対話システム用に特別にカスタマイズされた音声合成器を、株式会社東芝から御提供頂きました。

参考文献

- [1] Ancona, D. G. and Caldwell, D. F.: Beyond Task and Maintenance Defining External Functions in Groups, *Group & Or-*

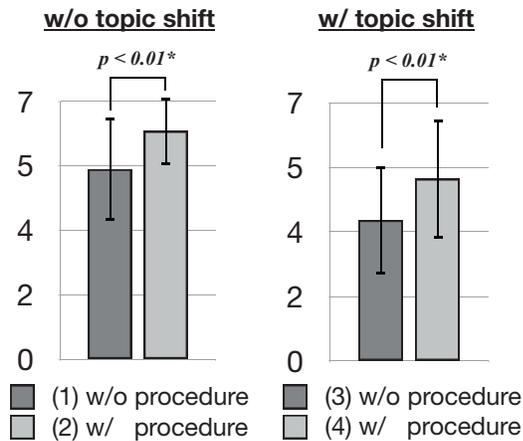


図5 実験 1-a (手続きの適切性) の結果

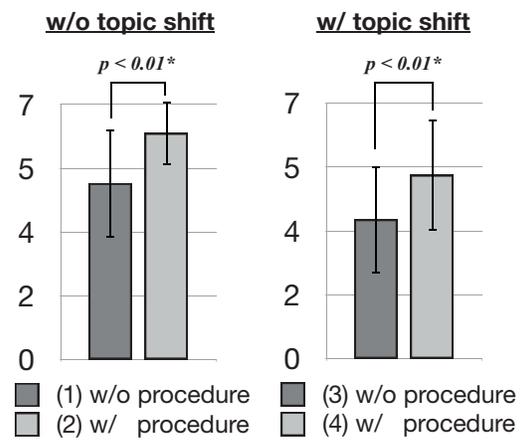


図6 実験 1-b (手続きを行ったことによる一体感) の結果

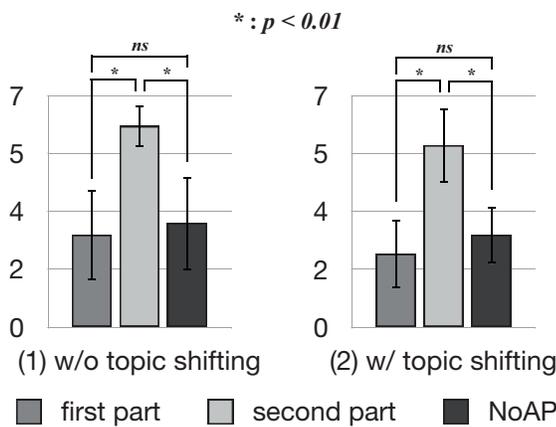


図9 実験 2 (手続き開始のタイミング) の結果

ganization Management, Vol. 13, No. 4, pp. 468–494 (1988).

[2] Bales, R. F.: Interaction process analysis, Cambridge, Mass (1950).

[3] Benne, K. D. and Sheats, P.: Functional roles of group members, *Journal of social issues*, Vol. 4, No. 2, pp. 41–49 (1948).

[4] Bohus, D. and Horvitz, E.: Facilitating multiparty dialog with gaze, gesture, and speech, *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, ACM, p. 5 (2010).

[5] Clark, H. H.: *Using language*, Vol. 4, Cambridge University Press Cambridge (1996).

[6] Dohsaka, K., Asai, R., Higashinaka, R., Minami, Y. and Maeda, E.: Effects of conversational agents on human communication in thought-evoking multi-party dialogues, *Proceedings of the SIGDIAL 2009 Conference: The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Association for Computational Linguistics, pp. 217–224 (2009).

[7] Goffman, E.: *Forms of talk*, University of Pennsylvania Press (1981).

[8] Kumar, R., Beuth, J. L. and Rosé, C. P.: Conversational Strategies that Support Idea Generation Productivity, in *Groups, 9th Intl. Conf. on Computer Supported Collaborative Learning, Hong Kong 160 and Rosé, 2010a) Rohit Kumar, Carolyn P. Rosé, 2010, Conversational Tutors with Rich Interactive Behaviors that support Collaborative Learning, Workshop on Opportunity*, Citeseer (2011).

[9] Mann, R. D.: A review of the relationships between personality and performance in small groups., *Psychological Bulletin*, Vol. 56, No. 4, p. 241 (1959).

[10] Matsuyama, Y., Tsuyoshi, T. and Kobayashi, T.: Conversation robot participating in group conversation, *IEICE transactions on information and systems*, Vol. 86, No. 1, pp. 26–36 (2003).

[11] Matsuyama, Y., Fujie, S., Taniyama, H. and Kobayashi, T.: Psychological Evaluation of a Group Communication Activation Robot in a Party Game, *Eleventh Annual Conference of the International Speech Communication Association* (2010).

[12] Matsuyama, Y., Taniyama, H., Fujie, S. and Kobayashi, T.: Designing communication activation system in group communication, *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on*, IEEE, pp. 629–634 (2008).

[13] Mudrack, P. E. and Farrell, G. M.: An examination of functional role behavior and its consequences for individuals in group settings, *Small Group Research*, Vol. 26, No. 4, pp. 542–571 (1995).

[14] Nakano, T., Fujie, S. and Kobayashi, T.: MONEA: message-oriented networked-robot architecture, *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, IEEE, pp. 194–199 (2006).

[15] Raux, A. and Eskenazi, M.: A finite-state turn-taking model for spoken dialog systems, *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, pp. 629–637 (2009).

[16] Sacks, H., Jefferson, G. and Schegloff, E. A.: *Lectures on conversation*, Vol. 1, Blackwell Oxford (1995).

[17] Stein, R. T. and Heller, T.: An empirical analysis of the correlations between leadership status and participation rates reported in the literature., *Journal of Personality and Social Psychology*, Vol. 37, No. 11, p. 1993 (1979).

[18] Williams, J. and Young, S.: Partially observable Markov decision processes for spoken dialog systems, *Computer Speech & Language*, Vol. 21, No. 2, pp. 393–422 (2007).