

卒業論文 2017年度 (平成 29年度)

会話の文脈と割り込み状況を考慮した
会話タスクスケジューリング手法 CACTS の実装と評価

指導教員

慶應義塾大学環境情報学部

高汐 一紀

徳田 英幸

村井 純

楠本 博之

中村 修

Osamu Nakamura

Rodney D. Van Meter III

植原 啓介

三次 仁

中澤 仁

武田 圭史

慶應義塾大学 環境情報学部

堀江 拓実

theramin@ht.sfc.keio.ac.jp

学部論文要旨 2017年度 (平成 29 年度)

会話の文脈と割り込み状況を考慮した 会話タスクスケジューリング手法 CACTS の実装と評価

論文要旨

近年、ロボットと人間が協調的に作業するための研究が多く行われている。一方、ロボットのインタフェースとしては音声認識が中心となっており、公共空間等では既に会話を開始しているロボットが会話の外から声を掛けられて割り込みを受ける、という可能性が考えられる。従来の会話中の割り込みに関する研究は、協調的なインタラクションの一部として解釈可能な割り込みについて扱うことが多く、それまでの会話の流れを無視した「会話外から割り込み」に関して研究したものは少ない。ロボットが割り込みに対応することを考えた時に、「会話外からの割り込み」を「会話を行うタスクの追加」とみなすと、割り込みの処理には、会話中の割り込みの検出、タスクのスケジューリング、およびタスク切り替え時の調整、の機能が必要となる。特に、タスクのスケジューリングは、割り込み主が常に適切なタイミングで割り込みを行うとは限らないため、重要である。本研究ではタスクのスケジューリングに注目し、会話の文脈と割り込み状況を考慮した会話タスクのスケジューリング手法として CACTS (Context Aware Conversation Task Scheduling) を提案する。CACTS は会話の重要度、長さ、割り込み主と割り込まれた側の人間関係、割り込み時の情動の 4 つの因子と会話シナリオの隣接ペアの情報によって会話のスケジューリングを行う。また、CACTS を用いたロボット会話システムを実装した。実装会話タスクの記述には、チャットボット用のマークアップ言語である AIML (Artificial Intelligence Markup Language) を隣接ペア単位で会話シナリオを記述するように独自に拡張した AIML-ap を用いた。実装したロボットを用いて CACTS の妥当性を評価を行なった。実験では、ロールプレイを通してロボットとの会話中に割り込まれる状況を体験してもらい、CACTS を用いた場合と割り込み主を信頼して常に最新のタスクを優先する場合の比較を行なった。実験の結果、CACTS は比較した手法に比べて好ましい印象を与えており、会話タスクのスケジューリングの手法として有効であることを明らかにした。また、CACTS の限界と CACTS 以外の割り込みの対応手法についての知見を得た。

キーワード

割り込み制御, 会話システム, ヒューマノイドロボット, H2R コミュニケーション, 隣接ペア

慶應義塾大学環境情報学部

堀江 拓実

Abstract of Bachelor's Thesis Academic Year 2017

Implementing and Evaluating CACTS : Context Aware Conversation Task Scheduling

Abstract

These days, many researchers have studied how to realize the human-robot cooperation. On the other hand, the voice recognition is often used as robot's user interface and there is the case that robot would be interrupted from outside conversation while the robot had been talking with another person. Traditional conversational interruption studies have aimed the interruptions as the cooperative interaction but few studies have aimed the interruption from outside conversation. In the case that a robot handles interruptions, assuming "interruption from outside conversation" as "addition of a task to conduct a conversation" and it is necessary to detect interruption during a conversation, task scheduling, and adjustment at task switching, for interruption handling. Especially, the task scheduling is important because interrupters also interrupt on inappropriate timing. In this research, I focus on the task scheduling and propose CACTS (Context Aware Conversation Task Scheduling) as a scheduling method of conversation tasks considering conversation context and interrupt situation. CACTS schedules conversations based on the information on the adjacent pairs of conversation scenario and the four factors of conversation importance, length, emotion at interruption and relation between the interrupter and the interrupted person. I also implemented a robot conversation system using CACTS. The system used AIML-ap, which is uniquely expanded AIML (Artificial Intelligence Markup Language), which is a markup language for chatbot, to describe the conversation scenario in adjacent pair units. I evaluated the CACTS with the implemented robot. In the laboratory work, subjects role-play and experiment interrupted situations and had to compare the impression of CACTS and the case when always giving priority to the latest task. Experimental results revealed that CACTS is more desirable than the baseline method, therefore, CACTS is an effective method for conversation task scheduling and we found the limitation of CACTS and other ideas of interruption handling.

Keywords

interruption handling; conversation system; humanoid robot; H2R communication; adjacency pairs

Keio University
Faculty of Environment and Information Studies
Takumi Horie

目次

| | | |
|-------|-----------------------|----|
| 第 1 章 | 序論 | 1 |
| 1.1 | 背景 | 1 |
| 1.2 | 関連研究 | 2 |
| 1.2.1 | 人とロボットの会話 | 2 |
| 1.2.2 | 割り込みに関する研究 | 3 |
| 1.2.3 | 現状のロボットの割り込み対応の問題点 | 3 |
| 1.3 | 本研究の目的 | 4 |
| 1.3.1 | 割り込み制御のフローと本研究の関係 | 4 |
| 1.4 | 本研究のアプローチ | 5 |
| 1.4.1 | 会話タスク中の中断可能なタイミング | 5 |
| 1.4.2 | タスクの優先順位づけ | 6 |
| 1.5 | 本論文の構成 | 6 |
| 第 2 章 | 会話の構造と中断可能なタイミング | 7 |
| 2.1 | 隣接ペアによる会話の構造化 | 7 |
| 2.1.1 | 隣接ペア | 7 |
| 2.1.2 | 隣接ペアの判定 | 8 |
| 2.1.3 | 隣接ペアの拡張 | 9 |
| 2.1.4 | 隣接ペアの会話システムへの応用 | 9 |
| 2.1.5 | 会話タスク中の中断可能なタイミング | 9 |
| 2.1.6 | まとめ | 11 |
| 2.2 | 隣接ペア構造における中断可能性の検証 | 11 |
| 2.2.1 | 評価方法 | 12 |
| 2.2.2 | 結果と考察 | 13 |
| 2.3 | まとめ | 15 |
| 第 3 章 | 中断可能性を考慮した会話スクリプト記述言語 | 16 |
| 3.1 | AIML | 16 |
| 3.1.1 | AIML の表現力 | 16 |
| 3.1.2 | AIML の問題点 | 18 |
| 3.2 | AIML-ap | 18 |
| 3.2.1 | AIML-ap の記述力 | 19 |

| | | |
|-------|------------------------------------|----|
| 3.3 | まとめ | 20 |
| 第 4 章 | 優先順位づけ因子を考慮した会話タスクスケジューリング手法 CACTS | 21 |
| 4.1 | 優先順位づけ因子の候補 | 21 |
| 4.2 | 各因子の影響力の分析 | 21 |
| 4.2.1 | 実験趣旨 | 21 |
| 4.2.2 | 因子の尺度と実験パターン | 21 |
| 4.2.3 | 実験方法 | 22 |
| 4.2.4 | 結果と考察 | 24 |
| 4.2.5 | まとめ | 28 |
| 4.3 | 会話タスクスケジューリング手法 CACTS | 29 |
| 4.3.1 | 会話タスクの切り替えの判断基準 | 29 |
| 4.3.2 | 会話タスクスケジューリングの流れ | 30 |
| 4.4 | まとめ | 33 |
| 第 5 章 | AIML-ap を用いた CACTS の実装と評価 | 34 |
| 5.1 | 実験システムの概要 | 34 |
| 5.1.1 | 会話タスクの内部表現 | 34 |
| 5.1.2 | 入出力インタフェース | 34 |
| 5.1.3 | 入力処理モジュール | 36 |
| 5.1.4 | 会話タスク実行モジュール | 37 |
| 5.1.5 | システムの動作の流れ | 37 |
| 5.2 | 実験 1 : CACTS の客観的な印象の評価 | 38 |
| 5.2.1 | 目的 | 38 |
| 5.2.2 | 実施内容 | 38 |
| 5.2.3 | 結果 | 38 |
| 5.2.4 | まとめ | 40 |
| 5.3 | 実験 2 : CACTS の割り込まれる側からの印象の評価 | 41 |
| 5.3.1 | 実験概要 | 41 |
| 5.3.2 | システムの修正 | 41 |
| 5.3.3 | 会話シナリオ | 41 |
| 5.3.4 | アンケート | 42 |
| 5.3.5 | 実験時の問題点 | 44 |
| 5.3.6 | 結果と考察 | 44 |
| 5.4 | まとめ | 52 |
| 第 6 章 | 結論 | 54 |
| 6.1 | 総括 | 54 |
| 6.2 | 今後の課題と展望 | 54 |
| 参考文献 | | 57 |

目次

| | | |
|------|---|----|
| 1.1 | クイズゲームロボット [1] | 4 |
| 1.2 | バーテンドーロボット [2] | 5 |
| 2.1 | Dyssuspendable Unit の決定手法に関する実験に用いた動画の一部 | 13 |
| 2.2 | 雑談をしているロボットへの割り込みにおける切り替えタイミングごとの印象評価 (* $p < 0.05$) | 15 |
| 2.3 | 問診をしているロボットへの割り込みにおける切り替えタイミングごとの印象評価 | 15 |
| 3.1 | AIML の例 | 17 |
| 3.2 | AIML-ap の例 | 18 |
| 3.3 | 挿入拡張の会話シナリオを追加した場合の AIML と AMLI-ap の記述例 | 20 |
| 4.1 | 割り込みシチュエーション再現 UI | 23 |
| 5.1 | システム構成図 | 35 |
| 5.2 | ビデオ内で会話タスクが実行されている時間 | 39 |
| 5.3 | ビデオ視聴後の各手法に対する印象評価の平均値 (**: $p < 0.01$) | 40 |
| 5.4 | 実験で使用されたロールの一例 | 42 |
| 5.5 | 3 回のインタラクション後の各手法に対する印象評価の平均値 (***: $p < 0.005$) | 44 |
| 5.6 | 3 回のインタラクション後の各手法に対する印象評価の平均値 (***: $p < 0.005$) | 45 |
| 5.7 | 実験者と被験者の人間関係別の 3 回のインタラクション後の各手法に対する印象評価の平均値 (\dagger : $p < 0.05$; 片側検定) | 45 |
| 5.8 | 実験者が人間関係で上位の時の各シチュエーションで好ましいと考える会話の切り替えタイミング | 47 |
| 5.9 | 実験者が人間関係で同格の時の各シチュエーションで好ましいと考える会話の切り替えタイミング | 47 |
| 5.10 | 実験者が人間関係で下位の時の各シチュエーションで好ましいと考える会話の切り替えタイミング | 48 |
| 5.11 | 実験者が人間関係で上位の時の各シチュエーションで好ましいと考える会話の切り替えタイミングと提案手法の一致率 | 48 |
| 5.12 | 実験者が人間関係で同格の時の各シチュエーションで好ましいと考える会話の切り替えタイミングと提案手法の一致率 | 49 |

| | | |
|------|--|----|
| 5.13 | 実験者が人間関係で下位の時の各シチュエーションで好ましいと考える会話の切り替えタイミングと提案手法の一致率 | 49 |
| 5.14 | 実験者が人間関係で上位だった場合の印象評価の結果 (†: $p < 0.05$; 片側検定) | 50 |
| 5.15 | 実験者が人間関係で同格だった場合の印象評価の結果 (*: $p < 0.05$, †: $p < 0.05$; 片側検定) | 50 |
| 5.16 | 実験者が人間関係で下位だった場合の印象評価の結果 (*: $p < 0.05$, ***: $p < 0.005$, †: $p < 0.05$; 片側検定) | 51 |
| 5.17 | 割り込んだ側が人間関係で上位だった場合の割り込んだ人への印象評価の結果 | 51 |
| 5.18 | 割り込んだ側が人間関係で同格だった場合の割り込んだ人への印象評価の結果 | 52 |
| 5.19 | 割り込んだ側が人間関係で下位だった場合の割り込んだ人への印象評価の結果 | 52 |

表目次

| | | |
|-----|-------------------------------------|----|
| 2.1 | 隣接ペアの例 [3] | 8 |
| 2.2 | 隣接ペアの拡張形の例 [4] | 10 |
| 2.3 | 雑談シナリオ | 13 |
| 2.4 | 問診シナリオ | 14 |
| 3.1 | Next タグの Type プロパティの種類 | 19 |
| 4.1 | 各因子候補の暫定尺度 | 22 |
| 4.2 | 因子の水準を揃える場合の値 | 22 |
| 4.3 | 説明変数と計算方法 | 24 |
| 4.4 | ロジスティック回帰分析の結果 | 25 |
| 4.5 | 各シチュエーションでの割り込み側が優先された割合 | 26 |
| 4.6 | CACTS における優先度の計算式と切り替えタイミングの閾値 | 29 |
| 5.1 | 会話の状態と音声認識の処理 | 36 |
| 5.2 | 実験で使った会話タスクの一覧 | 39 |
| 5.3 | 提案手法を用いた場合の各割り込みシチュエーションでの切り替えタイミング | 43 |
| 5.4 | 提案手法を用いた場合の各割り込みシチュエーションでの切り替えタイミング | 43 |

第 1 章

序論

本章では、本研究における背景および現状の問題点、そして本研究の目的、アプローチについて述べ、最後に本研究の構成を示す。

1.1 背景

先進国を中心に大きな社会問題となっているものの一つに少子高齢化などに伴う「労働者不足」がある。日本では 2016 年 10 月 1 日現在で 65 歳以上の高齢者人口が全人口に占める割合は 27.3% に達しており、少子化と高齢化の進行が予測されている [5]。少子化は労働人口の減少を、病気等のリスクの高い高齢者の増加は介護・医療分野の負担の増加を意味するため、介護・医療分野を中心に社会全体での労働力不足が予想される。そこで、近年では人間と協調して働くサービスロボットの活用が期待されている [6]。サービスロボットは、工場のような管理された環境ではなく、人間の社会生活空間の中での動作が求められる。したがって、予測困難で不確実な環境に適応しなければならない。実際に、Leite らはロボットを用いた長期的な実験について「健康・治療」「教育」「オフィス・公共空間」「家庭」の 4 つの実験フィールドに分類して分析を行い [7]、その中で、公共空間を対象とした研究の多くは、環境の不確実性に対処するために使い勝手の向上や外界への適用に特別な努力を行っていたことを発見している。ところで、サービスロボットのユーザインタフェースとしては、音声対話が多く用いられて来た。音声対話システムは臨床心理士を模した会話を行う ELIZA [8] や自然言語を理解して仮想空間内の積み木操作を行う SHRDLU [9] などのテキストベースの対話システムが源流となり、docomo のしゃべってコンシェル^{*1}や Apple 社の siri^{*2}などの一問一答型へと進化してきた [10]。また、近年では softbank 社の Pepper^{*3}や IBM の Watson Conversation Service^{*4}のように、シナリオベースの会話を行うシステムも多い。これらのシステムのほとんどは端末をユーザが占有して使うことを前提としており、途中から新たなユーザが現れるというような状況について考慮していないのだが、実際の人間同士のインタラクションでは、突然会話外からトイレの場所を尋ねられるというような事態が発生しうる。今後、人間と協調して動作することが期待されるロボットには、このような突然の割り込みという予測困難な事態においても適切な動作を行うことが求められると考えられる。

本研究では会話のみで完結するタスクのみを扱うことにする。理由として、現在普及しているコミュニケー

*1 https://www.nttdocomo.co.jp/service/shabette_concier/ (2017 年 12 月 17 日確認)

*2 <https://www.apple.com/jp/ios/siri/> (2017 年 12 月 17 日確認)

*3 <https://www.softbank.jp/robot/consumer/> (2017 年 12 月 17 日確認)

*4 [urlhttps://www.ibm.com/watson/jp-ja/developercloud/conversation.html](https://www.ibm.com/watson/jp-ja/developercloud/conversation.html) (2017 年 12 月 17 日確認)

ションロボットの多くは、非言語コミュニケーションのための最低限のアクチュエータしか持たないため、会話のみで完結するタスクだけでも十分に実用的と考えられるからである。会話のみで完結しないようなタスクの場合については、本論文の最後で改めて議論する。なお、これ以降本研究では必要がない限り、会話のみで完結するタスクとそのようなタスクのロボット内での内部表現を「会話タスク」、会話外からの割り込む相手と会話タスクを行うための割り込みを単に「割り込み」と呼ぶこととする。

1.2 関連研究

本論文で扱うのは、ヒューマンロボットインタラクション (Human Robot Interaction : HRI) と呼ばれる事象の一つである。本節では、まず会話を行う HRI について論点を整理し、次に、本論文のテーマとなる会話外からの割り込みについての関連研究を示す。

1.2.1 人とロボットの会話

人とロボットの会話についての研究が、その他の音声対話システムの研究と比較して特徴的な点として、ロボットは物理的な身体をもつことによる「コミュニケーションチャンネルの増加」と「空間的な広がり」が挙げられる。本節では、まず音声対話システム共通の話題について述べ、その後にロボットの会話インタラクション特有のテーマについて述べる。

音声対話システム全般のテーマ

音声対話システム研究の最初期の研究としては、MIT によって開発された街案内を行う音声対話システム VOYAGER[11] や、東芝によって開発された TOSBURG II[12] をあげることができる。この頃の特に重要な課題は、いかにロバストな音声認識を行うかという点であり、現在でも音声認識の精度の向上は重要なトピックとなっている [13]。また、音声対話システムとしての性能向上を目的として、誤りを含む音声認識結果を復元する手法 [14] や音声認識で使用する言語モデルの調整手法 [15] のような研究もなされている。また、システムによる音声の合成も重要なトピックであり、多数研究が行われている [16, 17]。一方、非言語的な部分に着目した研究としては藤江らの態度認識 [18] についての研究や大須らの話者交代の有無の推定についての研究 [19] などがある。

コミュニケーションチャンネルの増加

ロボットの身体性がもたらす特徴として、対話者に伝わる非言語的な情報の増加がある。ロボットの特徴的な非言語情報としては、視線 [20] や、相槌 [21]、モビリティを伴う振る舞い [22] などが挙げられる。これらを表現できるプラットフォームとして、robovie[23] が制作されており、多数の実験に用いられている [24, 25]。

空間的な広がり

身体性がもたらすもう一つの特徴として、インタラクションの空間的な広がりがあげられる。ロボットは、全方位からのインタラクションを受ける可能性をもっており、本研究が対象とする会話外からの割り込みの問題も、ロボットの身体性が生む空間的な広がりによって生じていると言える。実際の研究例としては、松坂らによる多人数会話に参加するロボット [26] や、秋谷らによる介護施設で点在する人と円滑にコミュニケーションを開始するロボット [27]、1章で触れた Petrick らのバーテンダーロボット [2] や Bohus らの多人数でクイ

ズゲームを行うロボット [1] など様々な例がある。多人数の会話については Goffman による聞き手の分析が参照されることが多い。Goffman は聞き手の役割を、直接話の対象となっている「受け手」、直接の話の対象となっていないが会話への参与が認められている「傍参与者」、偶然聞いてしまった「傍観者」、こっそりと聞き手となっている「盗み聞き者」に区分しており^{*5}[29]、松坂らの研究では、グループ会話におけるそれぞれの役割を推定し次の発話者の予測などに利用している。

1.2.2 割り込みに関する研究

聞き手による割り込み

多くの場合、会話中の割り込みに関する研究は対話中に聞き手が行う割り込みを対象にしてきた。Sacks らは、まず会話の進行はどちらか一人が話し手となり、話し手が交互に入れ替わることで成立するという原則を定め、この原則に反して相手の発話中に重ねて発話をおこなうことを「割り込み」と定めた [30]。その発生タイミングや原因、性質については様々な研究がなされている [31, 32]。堂坂らは聞き手からの割り込みをハンドリングする手法を検討している [33]。しかしながら、これらの研究が対象とする「割り込み」は、本研究が対象とする会話タスクを要求する会話外からの割り込みとは性質が異なるものである。

会話外からの割り込み

会話外からの割り込みは日常生活で自然に発生しており、会話外からの割り込みを行う理由は 2 つに分類できると考えられる。一つは現在行われている会話への参加のためであり、もう一つの理由は、高梨ら [34] が観察した例のように割り込み主が割り込む相手と行いたいタスクを持つ場合である。前者の会話中の新たな話者の参入については、多人数ロボット対話の分野においても研究が行われている [35, 36]。一方、後者の場合を扱った研究は少ない。Schegloff は、本研究のような割り込みを「挿入の割り込み」と呼んで聞き手からの割り込みと区別し、割り込みが発生した際の優先順位の評価基準として、「地位」と割り込みが要求するインタラクションについての「時間的評価」を挙げている [37]。さらに「時間的評価」には『割り込みの広がり (etensiveness) の適切な「比率」(真剣さ、重要性、優先性などの比較考量に基づいて、それ自体変化すると思われる比率) への注意が存在する』と考察している。さらに、会話外からの割り込みが発生した時点で割り込み主は自身が優先権を持つと認識しており、割り込まれた側も割り込み主の判断を理解した上で対応するので原則的に割り込みが起きると新たに要求される会話の方が優先されるだろうと考察している [37]。

しかし、高梨らが観察したように [34]、実際には割り込み主が十分に思慮して判断したとは思われない事例も多く、常に割り込み主の判断を信用するのは妥当とは言えないため、ロボットによる会話のスケジューリングは必要と考えられる。

1.2.3 現状のロボットの割り込み対応の問題点

会話外からの割り込みとは、先ほどのトイレの場所を尋ねる例に代表されるように、それまで実行されてきた会話タスクに並行して異なる会話タスクを行うことを要求することと言える。HRI において、会話タスクに限らず複数のタスクの実行が要求されるような場面を扱った研究例としては、Petrick らによる図 1.2 のパートナーロボット [2] がある。このロボットは任意のタイミングで客がやってきて、同時に複数人の接客を行わなければならない場合でも、行動をスケジューリングして実行することができる。また、Bohus らは図

^{*5} 用語の訳は [28] を参照した

1.1のような多人数でクイズゲームを行うロボットを構築した [1]。このロボットは途中で新たな参加者が現れた場合、クイズゲームの進行度と参加者の回答状況に応じて対応を変更する。しかしながら、これらのロボットの行動ルールはタスクの内容に強く依存しており、汎用的に利用するのは困難な状態であり、複数種類のタスクをスケジューリングすることは困難だった。



図 1.1: クイズゲームロボット [1]

1.3 本研究の目的

本研究の目的は、会話タスクの詳細に左右されない、汎用的な会話タスクスケジューリング手法を提案することである。

1.3.1 割り込み制御のフローと本研究の関係

先行研究のタスクスケジューリング手法を観察すると、人の接近などをトリガーとして、タスクを追加し、スケジュールを遂行する、という一連の流れが存在する。これを割り込みに置き換えて、複数のタスクに対応させることを考えた場合、割り込み制御のフローは次のようになると考えられる。

1. 割り込みの発見
2. 割り込みの意図に合わせた会話タスクの生成
3. 新たに生成された会話タスクのスケジューリング
4. 会話タスクの円滑な切り替え

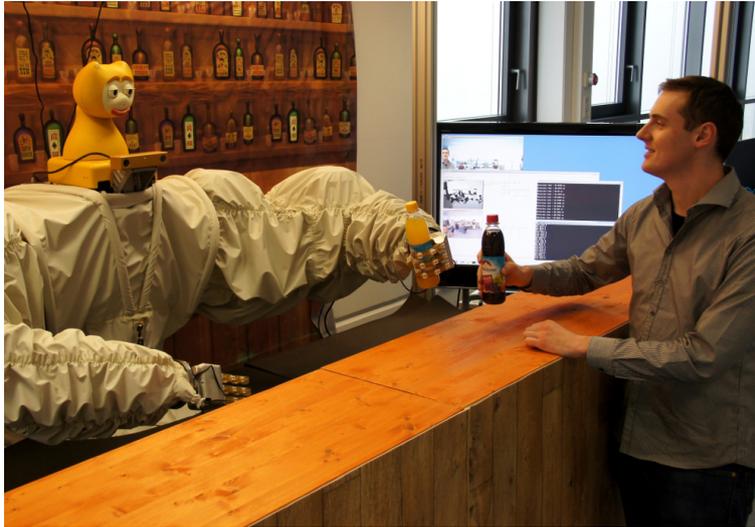


図 1.2: バーテンダーロボット [2]

ロボットが割り込みの制御を行うには、まず割り込みを発見することが必要であり、さらにその意図を把握してロボットへの要求を理解する必要がある。そして、要求と合致する会話タスクを生成すると、今度は、現在実行されている会話タスクと現在実行されていない会話タスク、そして新たに生成された会話タスクを合わせて、どの時点でどの会話タスクを実行すべきかを決定する必要がある。そして、実際に会話タスクを切り替える際には、会話相手に違和感や不快感を与えないような円滑な会話タスクの切り替えを行う必要がある。本研究は、このフローの中でも「新たに生成された会話タスクのスケジューリング」を対象としている。

1.4 本研究のアプローチ

ロボットがユーザを不当に待たせることは好ましくないことから、ロボットにはレイテンシを小さくすることが求められる。従来の研究のように、スケジューリングすべきタスクの種類が限定されているならばデザイナーによるタイミングの経験的な調整も容易だが、タスクの種類が増えていったときにスケールしない。ラウンドロビンのような手法でロボットが会話タスクを切り替えた場合、話し相手が短い期間で切り替わることとなる。計算機内のプロセスと異なり、人間はロボットとの会話タスク以外にも実行すべきで実行可能なタスクが存在する可能性があるため、待つ側の待ち時間が断片化するラウンドロビンは好ましくない。本研究では、低レイテンシと悪印象を与えない会話タスクの切り替えを両立するスケラブルな手法として、状況に応じた会話タスクの優先順位づけを行い、キリが悪い区間を避けて会話タスクの切り替えを行う会話タスクスケジューリング手法 CACTS を提案する。本節では、CACTS における会話タスクの優先順位づけと会話のキリの良いタイミングの判断のアプローチについて述べる。

1.4.1 会話タスク中の中断可能なタイミング

本研究では、会話タスクが中断可能なタイミングであるかの判断材料として、会話構造の一種である隣接ペア [3] という概念を用いる。詳細は 3 章で述べるが、隣接ペアは、時間的に隣り合った二つの発話同士の関係性に注目した概念であり、この隣接ペアの途中のうちに会話が中断されることが敬遠されていることを示唆す

る研究 [38, 39] があり，隣接ペアには中断可能なタイミングとの関連性がある．

1.4.2 タスクの優先順位づけ

待ち時間の断片化が起こらず低レイテンシを実現する手法として，リアルタイム OS などで採用されるような優先度ベースのスケジューリングを採用する．Schegloff の考察した会話の優先度を決定する要素のうち，真剣性の直接の取得は困難であるが表情が笑顔の場合より真顔の方が真剣、というような表情をベースにした議論は可能であると考えられる．そこで，本論文の 5 章で以下の因子が実際に影響力を持つかについて議論を行う．

1. 聞き手から見た割り込み主との人間関係
2. 会話タスクの長さ
3. 割り込み時の会話タスク参加者の感情
4. 会話の重要度

1.5 本論文の構成

本研究の構成は以下のようになっている．2 章では会話構造と会話の中断可能なタイミングの関係性を明らかにするための検証実験について述べ，3 章では 2 章の議論に基づいて独自に定義した，会話スクリプト記述言語 AIML-ap について述べる．4 章では会話タスクの優先度因子の影響力の比較についての実験を行った上で，会話タスクのスケジューリング手法 CACTS の詳細について述べる．5 章では，まず CACTS の検証するためのロボットシステムの実装について述べ，CACTS を用いた場合と常に最新の会話タスクを最優先するスケジューリングを行う場合とを比較する実験の詳細について述べる．1 つ目の実験ではロボットが会話を切り替える様子をビデオを通して客観的な立場で判断してもらった．2 つ目の実験では，実際に被験者に割り込まれる側になってもらってロボットの対応方法によって受ける印象の違いを実験で評価し，CACTS の有用性について議論する．6 章で本研究の総括を行う．

第 2 章

会話の構造と中断可能なタイミング

本章では、従来研究によって示されている会話の構造について示した上で、その構造上の切れ目が会話タスクの中断可能なタイミングであるかについて検証する実験について説明し、結果を議論する。

2.1 隣接ペアによる会話の構造化

人間同士の会話について分析する手法として、Sacks や Schegloff らによって開発された会話分析 [?] と呼ばれる手法がある。特に、会話分析で用いられる隣接ペアという概念は会話の構造化を行う上で効果的である。本節では隣接ペアに関する知見について述べる。

2.1.1 隣接ペア

会話において、一方の質問に対しもう一方の話者がすぐに質問に対する応答を返す、という流れは非常に自然である。Schegloff は電話上での会話を書き起こしたものについて分析を行い、先ほど挙げた「質問－応答」のペアや「提案－承認/拒否」、「挨拶－挨拶」のような発話の対が隣り合う形で繰り返し会話中に現れることを発見し、このような発話対のことを会話インタラクションの最小単位として「隣接ペア」と名付け、次のような原則を守る親密な関係性を持った発話対であると定義した [3]。

1. 二つの発話から成る
2. 二つの発話が隣接した位置に配置されている
3. 各々の発話をそれぞれ別々の話し手が生成する
4. 二つの発話の間には相対的な順序が存在する
5. 一つ目の発話はどのような二つ目の発話を選択すべきかに影響を与える

一般に隣接ペアの先行して発話される部分を第 1 部分、続いて発話される部分を第 2 部分と呼ぶ。例えば「質問－応答」の隣接ペアであれば、質問が第 1 部分、応答が第 2 部分である。5 番目の原則は、第 1 部分が発話された時点で対となる第 2 部分が発話されることが強く期待されるようになることを意味している。例えば現在の時間についての質問をした場合、次の話者は現在の時間についての応答以外をすることは稀で、突然昨日の夕飯について話し出すようなことはないという原則を表している。表 2.1 は Schegloff が実際に会話分析を行なったスクリプトの 1 つで、電話越しの会話の冒頭部分である。この例では、B の最初の発話「おはよう」と次の「ああ、おはよう」が「挨拶－挨拶」の隣接ペア、最後の「起こしちゃったんじゃない」と「いいえ、

そうじゃないわ、新聞を読んだの…」という発話対が「質問-応答」の隣接ペアとして成立している。

表 2.1: 隣接ペアの例 [3]

| 話者 | 発話 |
|----|----------------------|
| A | もしもし |
| B | おはよう。 |
| A | あぁ、おはよう。どう、元気、フフウ |
| B | リサね |
| B | 元気よ、起こしちゃったんじゃない |
| A | いいえ、そうじゃないわ、新聞を読んだの… |

2.1.2 隣接ペアの判定

一方、Schegloff の定義では隣接ペアであるためには第 1 部分が第 2 部分になる発話を制約している必要があるため、意見や情報提供のような発話は受け手の発話に制約を課さないことから第 1 部分とはなり得ない。しかしながら、隣接ペアの認定に際していくつかのケースで隣接ペアかどうかの判定が難しい事例が存在している。実際に会話コーパスに対して隣接ペアに基づくタグ付けを行った研究 [40] で、アノテーションを行なった 4 人の間でタグ付けが一致せず、タグ付けの判定が困難だった点について報告している。ここではその中でも「相槌」と「発話の複合的単位」について述べる。

普通、発話は誰が話すかという会話の主導権を持っている人が行う。従って、会話の主導権が移譲されずの間に挟まるだけの相槌では、一般には 1 つの発話としては扱われないため隣接ペアの要素とはならない。しかしながら、[40] では相槌が存在しなければ不自然さが生じるような事例として、「か」や「ね」のようなあいづちや同意を求めるニュアンスを持つ終助詞に続く相槌や、長い沈黙の際に発せられる相槌について言及している。反語表現の「か」の例として、海外旅行に関して前提条件を確認する「わりとこう最近若いうちにいっぱい色んなところ行くじゃないですか」という発話を挙げている。この場面では「行くじゃないですか」に続いて「はいはい」という相槌が行われており、「行くじゃないですか」の反語表現が次に相槌が続くように制約していると捉えて、[40] ではこの例を隣接ペアに認定している。一方「ね」の例としては、バラバラになった 4 コマ漫画のコマの並び順を決定するタスクにおける、話し手が持っている漫画のコマについて説明した、「でねなんか、ひと一方が眼鏡を外して、なんかハンカチで拭こうしてるのね」という発話を挙げている。この発話では直後に「うん」という相槌が後続しており、最後の「うん」がないと違和感が感じると判断したアノテータがいた。しかし「ね」が使われる場面で必ずしも相槌が続くとは限らず、「か」の終助詞と比べると頻度が低いため、「ね」を用いる場合は従来の会話分析では隣接ペアとはみなされてこなかった。さらに、芸能人をギャラの順で並べるタスク中の会話で、「じゃそれを蟹瀬にして」という発話がなされて約 2.5 秒後に「はい」と続くという場面があり、この「はい」という相槌が同意表現としての発話なのか、あるいは会話の続きを促すための相槌表現なのかという点についても意見が割れていたことを報告している。

また、インタビューを行う事例において、質問に対する応答が非常に長く、対応する発話が複数存在するようなケースがあったことを報告しており、[40] では、複数の発話のうち最初の 1 つだけを第 2 部分とみなしたり複数の発話それぞれを第 2 部分と認定するよりも、これらの複数の発話を 1 つの複合的単位として第 2 部分と捉えるべきではないかと述べている。

2.1.3 隣接ペアの拡張

上記に加えて schegloff は隣接ペア同士にも強い結びつきがあるものを発見し、隣接ペアの拡張形を定義した。schegloff の著作 [4] に従うと、以下のような拡張形がある。

前方拡張 メインとなる勧誘や要求の隣接ペアに先立って、それが行われることが適切かを準備的に確かめる隣接ペアが置かれる場合をさす。不適切であることがわかった場合はメインとなる隣接ペアは実行されない。

挿入拡張 第一部分と第二部分の間に別の隣接ペアが挟み込まれる場合をさす。「質問-応答」の隣接ペアにおいて質問意図が理解できないなど、現在の状態では第二部分を発話できない場合に、確認のために発生する。この場合元々の隣接ペアは隣り合わなくなるがこの場合は隣接ペアとみなす。

後方拡張 直前の隣接ペアの第二部分に関して追加の隣接ペアが交わされる場合をさす。

連鎖終結の第三部分 隣接ペアの第二部分を受けて、「oh」や「okay」などの聞き取ったことを示す語や感想が述べられた発話を指し、元々の隣接ペアの第三部分とみなす。

拡張形の例を表 2.2 に示した。発話の最後に加えた記号は隣接ペアとしての構成部分に対応しており、 F が第一部分、 S が第二部分、 T が第三部分を意味し、小文字のものは拡張によって付け加えられた隣接ペアであることを表している。前方拡張の例では、 A が飲みものが欲しいかを聞く前に、まず飲み物を欲しがるといった状況を訪ねている。挿入拡張の例では A の最初の質問の一部が聞き取れなかったため、確認を行なう隣接ペアが挟み込んだ形式になっている。後方拡張の例では、 B の回答が満足ではなかったため、詳しい説明を求める隣接ペアが続いている。最後の連鎖終結の第三部分の例では、調子を尋ねる質問の回答に対して感想を A が返している。

このように隣接ペアおよび、隣接ペアの拡張形とみなすことのできる発話の間には特別な結びつきを見ることが出来る。

2.1.4 隣接ペアの会話システムへの応用

隣接ペアの概念は第 1 部分が発話された時点で第 2 部分を制約するという性質から、会話システムの応答生成に部分観測マルコフ決定過程を用いる際の学習データとして用いられたりする [42]。一方、隣接ペアの概念を用いた会話の構造化に注目して会話システムに応用した例として、アンドロイドの相槌などのフィラー生成タイミングを決定するために会話中の隣接ペアの構造を用いた研究 [43] や、秋葉らによる、参加者の発話量の調整を行うファシリテーションロボットがある [38]。秋葉らのロボットは Goffman の提案した参与構造 [29] における「傍参与者」のルールとして会話に関わり、発言量の少ない参加者が存在する場合、ロボットが会話の主導権を奪取してその参加者に主導権を移譲することで参加者間の発話量が同程度になるように調整する。秋葉らがこのロボットを被験者に使用させたところ、ロボットの主導権奪取のための介入のタイミングは、隣接ペアの途中より隣接ペアが完成直後の方がロボットに対して良い印象を抱くことを発見した。

2.1.5 会話タスク中の中断可能なタイミング

ロボットが割り込みを受けた際、新しいタスクの方が優先すべきと判断された場合でも即時の切り替えが必要とされないのであれば、現在の会話タスクの”キリがいい”ところまで進めてから会話を切り替えるという挙動

表 2.2: 隣接ペアの拡張形の例 [4]

| 名前 | 例 |
|-----------|---|
| 前方拡張 | (伝による [4] の例 (4.01) の和訳 [41]) A : もしもし B : やあ A : やあ B : 何してる? /f A : これとって /s B : 飲みたい? /F A : うん . /S B : オーケー /T |
| 挿入拡張 | ([4] の例 (6.01) の和訳) A : ミッシズ・ケリーにあったのは昨晚が初めて? /F B : 誰に会ったって? /f A : ミッシズ・ケリー . /s B : そうだよ . /S |
| 後方拡張 | ([4] の例 (7.37) の冒頭の和訳) A : それで、君はどこで働いているの? /F B : えーっとアムファカンパニーってとこ . /S A : それどこ? /f B : アムファカンパニー . 持ち株会社なんだよ . /s |
| 連鎖終結の第三部分 | ([4] の例 (7.09) の冒頭の和訳) A : お元気? /f B : まあ多分生きてるかな . /s A : そいつはいいや /t |

は自然だと考えられる。この”キリがいい”部分と隣接ペアとの関連性を示唆する研究として、前節で述べた秋葉らの研究の他に、Shyrovkov らによる実験がある [39]。Shyrovkov らの実験では、まず被験者はドライビングシミュレータの助手席に乗ってもらい、何も無い時には運転手に道案内を行ってもらう。そして、道案内の最中、被験者に対して画像が提示されるので、被験者は画像と一緒に表示された緊急度を参考にして好きなタイミングで運転手に提示された画像の説明を行ってもらおうよう指示した。結果、画像が提示されたタイミングが隣接ペアを構成する発話の途中である場合、被験者は隣接ペアの終了まで提示された画像についての説明を行うことを待つ傾向にあることがわかった。このように、従来の研究では、隣接ペアの途中で現在の会話を中断するような行為は敬遠されていることを示している。ただし、秋葉らと Shyrovkov らのどちらの研究でも隣接ペアの拡張形については言及していない。

2.1.6 まとめ

本研究で扱う割り込みとは、Goffman の提案した参与構造の枠組みで捉えた場合、参与が認められていない者が承認の有無を無視してそれまでのトピックとは異なる話題の会話を行うことを要求する行為であると言える。秋葉らのような傍参与者による介入に関する研究も、参与の承認の有無の差はあれど、会話の中心である話者と受け手の外から会話に介入していくという点で本研究における割り込みとの類似性を認めることができる。

2.2 隣接ペア構造における中断可能性の検証

前節で述べたように、隣接ペアの途中で現在の会話を中断するような行為は敬遠されていることから、会話外からの割り込みに対する会話タスクの切り替えも、隣接ペアの途中で行うことを避けて隣接ペアの完成まで保留する可能性が考えられる。一方従来研究では、隣接ペアの拡張形の途中で現在の会話タスクを中断する行為が敬遠されるといったことは報告されていないものの、隣接ペアの拡張形は、1 つ目の隣接ペアが次の隣接ペアを予期させる、という隣接ペアの第 1 部分と第 2 部分の関係性に近い結びつきを持つことから、隣接ペアの拡張形の途中で現在の会話タスクを中断するような行為も同様に敬遠される可能性が考えられる。そこで、会話外からの割り込みを受けた際の会話タスクの切り替えを、隣接ペアあるいは隣接ペアの拡張形の完成まで保留すべきかについて明らかにするための検証実験を行った。

実験には、ロボットが会話外から割り込みを受けるシチュエーションを再現した動画を用いた。動画は、割り込む会話タスクと割り込まれる会話タスクの内容は同一だが、割り込みを受けた後にロボットが会話タスクを切り替えるタイミングだけ異なるものを複数用意し、それぞれの動画において被験者がロボットに対して受ける印象を調査する。本実験では、隣接ペアや隣接ペアの拡張を考慮する必要はなく、すぐに会話タスクを切り替えて良いのか、あるいは隣接ペアの拡張形は無視しても、隣接ペアの途中での会話タスクの中断は避けの方が良いのか、さもなければ隣接ペアの拡張形の途中での会話タスクの中断も避けるべきなのかを知りたいので、動画中の会話の切り替えタイミングとして次の 3 つを採用した。

1. 割り込まれてすぐ
2. 隣接ペアの拡張形の途中であるものの、割り込まれた時に途中だった隣接ペアは完成したとき
3. 隣接ペアの拡張形が完成したとき

割り込まれてすぐの場合の印象が他よりも高い場合、会話タスクの切り替えを保留する必要はないと言える。また、「隣接ペアの拡張形の途中であるものの、割り込まれた時に途中だった隣接ペアは完成したとき」が最も印象が良い場合には、隣接ペアの拡張形の途中であるかは会話タスクの切り替えが敬遠される理由とはならず、単一の隣接ペアの完成まで会話タスクの切り替えを保留すればいいとみなすことができる。最後に、「隣接ペアの拡張形が完成したとき」が最も印象が良い場合には、隣接ペアの拡張形が完成するまでは会話タスクの切り替えを保留すべきと言える。

一方、割り込みのシチュエーションについては、割り込み側が優先されやすいが、優先されやすさの度合いが異なると考えられる 2 つシチュエーションを用意し、優先されやすさの度合いが、会話タスクの切り替えタイミングの好ましさに影響を与えているかについても調査するため、具体的には次の二つのシチュエーションを採用した。

1. 問診中のロボットが予定確認のために割り込まれる
2. 雑談中のロボットが予定確認のために割り込まれる

1つ目の健康診断に割り込むシチュエーションでは、予定確認というシンプルで短く完了すると考えられるという観点から、割り込み側が優先されると考えられる。一方、雑談に割り込むシチュエーションでは、予定確認というシンプルで短く完了すると考えられるという観点に加え、割り込み側のみ明確な目的を持っているという観点からも、割り込み側が優先されると考えられるため、2つ目のシチュエーションのほうが、より、優先されやすさの度合いが高いと考えられる。

雑談の会話タスクの会話進行は表 2.3 に、問診の会話タスクの会話進行は表 2.4 に示した。両端に横線がついた文は、割り込みおよび会話タスクの切り替えのタイミングについて表す。また、発話の最後につけた記号はシナリオ内の隣接ペアおよび隣接ペアの拡張系の構造を表しており、まず、 $F(f)$ はその発話が第 1 部分であることを、 $S(s)$ はその発話が第 2 部分であることを、そして $T(t)$ はその発話が第 3 部分であることを表している。また、大文字の F, S, T の場合にはそれらが隣接ペアの拡張形においてメインとなる隣接ペアの構成要素であることを示しており、小文字の f, s, t である場合にはそれらが、隣接ペアの拡張系のうち、拡張によって追加された隣接ペアの構成要素であることを示している。また、下付き文字はその隣接ペアの拡張系の種類を表しており、 PRE は前方拡張、 $INSERT$ が挿入拡張、 $AFTER$ は後方拡張を意味している。動画内での会話タスクへの割り込みタイミングは、隣接ペアの拡張形が完成したときに会話タスクを切り替える場合の切り替えまでの時間になるべく長くなり、直近の隣接ペアが完成した場合と差別化できるようなタイミングとして、どちらも隣接ペアの拡張形の中の 2 つ目の隣接ペアの途中で割り込むようにした。具体的には、次のタイミングである。

雑談 「いつもいく場所とかあるんですか？」の直後
健康診断タスク 「どうかされたんですか？」の直後

最後に、動画の一部を図 2.1 に示す。撮影は図 2.1 のような室内で行われた。ユーザ役は当研究室の学生二人で普段からロボットとの会話経験は豊富である。一人目のユーザは動画開始時からロボットと会話タスクを行っており、割り込み役の学生は割り込みのタイミングになるとロボットから見て右側、動画では奥側にやってきて割り込みを行う。まず、ロボットは割り込みが起きると、割り込んできたユーザを一度見て「はい」という相槌を打って割り込みを認識したことを明示する。割り込まれてすぐに会話タスクを切り替える場合には、そのまま割り込んできたユーザと会話を行い、会話終了後元のユーザとの会話に戻る。一方、割り込んできたユーザに一度待ってもらった場合には「少々お待ちください」と発話した後、元のユーザとの会話を実行し、隣接ペアや隣接ペアの拡張形が完成されたタイミングで、割り込んできたユーザに再び向き直って会話をおこなう。

2.2.1 評価方法

web ブラウザ上で、7 段階の SD 法と自由記述としてロボットの会話中の振る舞いについての感想、および気付いた点について調査した。印象評価の形容詞対には以下の 5 つを用いた。

- 「失礼な - 礼儀正しい」
- 「ぎこちない - 自然な」
- 「融通の利かない - 柔軟な」

表 2.3: 雑談シナリオ

| | |
|---|---|
| R | そういえば、よく散歩に行きますよね。 / f_{PRE} |
| A | うん、よく行くね。 / s_{PRE} |
| R | いつも行く場所とかあるんですか? / F —割り込みの発生 / 割り込まれてすぐの切り替えタイミング— |
| A | うーん、川とかかな。 / S |
| R | へー、そうなんですネ。 / T —割り込まれた際に途中だった隣接ペアの完成まで保留した場合の切り替えタイミング— |
| R | 北のほうの川ですか? / f_{post1} |
| A | うん、そうだよ。 / s_{post1} |
| R | いいですね。 / t_{post1} 北の川は綺麗なところですよネ。 / f_{post2} |
| A | うん、綺麗だよネ。 / s_{post2} —割り込まれた際に途中だった隣接ペアの拡張形の完成まで保留した場合の切り替えタイミング— |

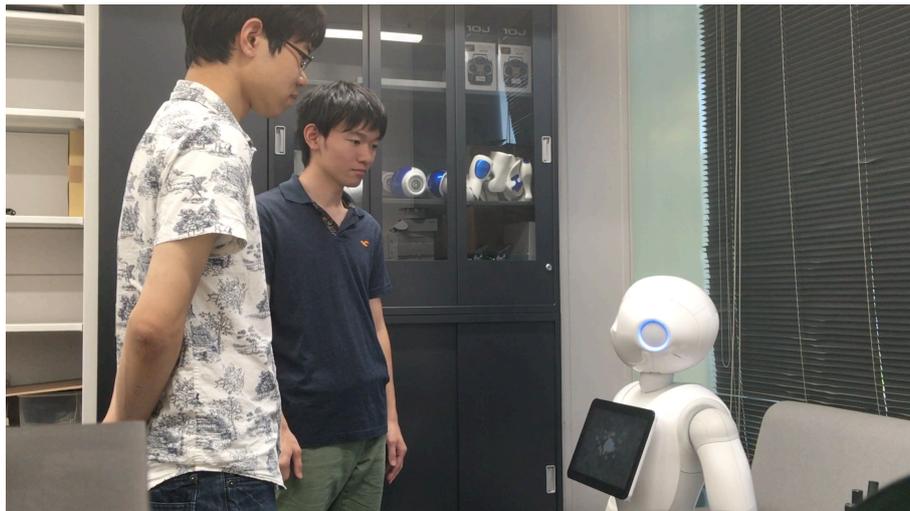


図 2.1: Dyssuspendable Unit の決定手法に関する実験に用いた動画の一部

- 「愚かな - 賢い」
- 「適切な - 不適切な」

2.2.2 結果と考察

被験者は 9 名でいずれも大学生、男性が 7 名、女性が 2 名が参加した。平均年齢は 20.6 歳であった。ただし、うち男子 1 名の評価データの記入にミスがあったため、印象の評価の項目では除外している。

印象評価の結果は、雑談をしているロボットに割り込む場合が図 2.2、問診をしているロボットに割り込む場合が図 2.3 のようになった。問診をしているロボットに割り込みを行うシナリオにおいて、「それは心配で

表 2.4: 問診シナリオ

| | |
|---|--|
| R | おはようございます． / F |
| A | おはよう． / S |
| R | 今日の体調はいかがですか？ / F |
| A | あんまり良くないかな． / S |
| R | どうかされたんですか． / f_{post1} |
| | —割り込みの発生 / 割り込まれてすぐの切り替えタイミング— |
| A | 熱っぽいんだよね． / s_{post1} |
| R | それは心配ですね． / t_{post1} |
| | —割り込まれた際に途中だった隣接ペアの完成まで保留した場合の切り替えタイミング— |
| | ほかに悪いところがありますか？ / f_{post2} |
| A | 鼻水もなんだよね． / i_{insert} |
| R | 鼻水もですか． / s_{insert} |
| A | うん． / t_{insert} |
| | 他にはないかな． / s_{post2} |
| R | わかりました．あとでしっかり検査しましょう． / t_{post2} |
| | —割り込まれた際に途中だった隣接ペアの拡張形の完成まで保留した場合の切り替えタイミング— |

すね」という発話の直後（隣接ペアの完成直後）に会話を切り替えた場合の自然さの値が他の切り替えタイミングの場合と比べて有意に低い値になった。また、有意差は無いものの、礼儀正しさ以外の項目では隣接ペアの基本形の完成まで会話タスクの切り替えを保留した場合の評価は他のタイミングの場合と比べて同じかより低い評価になっている。これは、隣接ペアの拡張形の途中での会話の切り替えが敬遠されていることを表していると言える。コメントでも隣接ペアの基本形の完成まで会話タスクの切り替えを保留した場合については「中途半端なタイミングで話を切り替えている」というコメントや「話し始めたのにまた途中で止めることもあまりよろしくない」などのコメントがあり、被験者が切り替えタイミングとしては中途半端であるという印象を抱いていることを示している。

一方、すぐに会話タスクを切り替えた場合と隣接ペアの拡張形の完成まで保留した場合とでは印象に有意な差はなかったが、雑談をしているロボットに割り込む場合では、すぐに会話タスクを切り替えた場合の方が隣接ペアの拡張形の完成まで保留した場合よりも印象評価の値が良かった項目が3つあったのに対し、問診をしているロボットに割り込んだ場合では、全ての項目で印象評価の値が隣接ペアの拡張形の完成まで保留した場合の方が良い結果となった。このような結果となった理由として、優先度の差の大きさの違いが考えられる。雑談をしているロボットに割り込む場合には、会話の長さの差以外にも明確な目的を持っているかという点でも差がある。そのため、問診をしているロボットに割り込む場合に比べて、割り込む側と割り込まれる側の会話タスクの優先度の差が大きくなり、雑談をしているロボットに割り込む場合の方がより早く会話を切り替える方が望ましいと被験者が判断しているのではないかと考えられる。実際に、雑談をしているロボットに割り込む場合では自由記述で「スケジュールの方の話を優先すべき」や「ベースの会話が雑談だと、スケジュール確認の方が重要に思えるので雑談を続けることが不自然に思いました」というコメントがあり、被験者側にもはっきりと優先度の差が意識されていたことがうかがえるコメントが見られた。

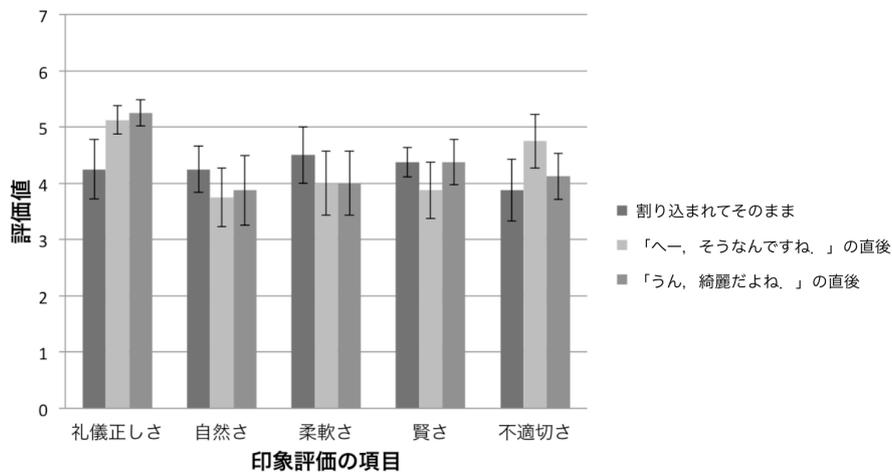


図 2.2: 雑談をしているロボットへの割り込みにおける切り替えタイミングごとの印象評価 (* p < 0.05)

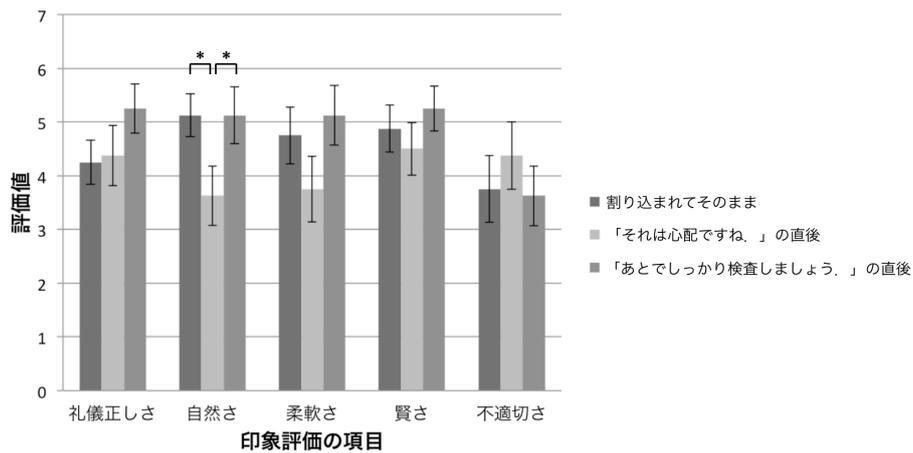


図 2.3: 問診をしているロボットへの割り込みにおける切り替えタイミングごとの印象評価

2.3 まとめ

本章では、会話分析における会話インタラクションの最小単位である隣接ペアとその拡張形による会話の構造化について述べ、それらの構造を考慮した会話タスクの切り替えを行うべきかについての検証実験の詳細を述べた。検証実験の結果、隣接ペアの拡張形の途中での会話の切り替えは望ましくないことがわかった。また、割り込む会話タスクと割り込まれる会話タスクの差が大きい場合には、割り込まれてすぐに会話を切り替えた方が望ましいと考えられ、会話システムは会話中の隣接ペアおよびその拡張形の情報を認識できるようにすることが、割り込み制御において重要であることがわかった。

第 3 章

中断可能性を考慮した会話スクリプト記述言語

前章では、隣接ペアの拡張形の途中での会話の切り替えは望ましくないことがわかった。しかし、既存の会話シナリオ記述言語では隣接ペアの拡張形までを記述したものは存在しないため、旧来の会話シナリオ記述言語を用いた場合には、ロボットシステムが切り替え可能なタイミングを決定することは容易ではない。そこで、本章では会話システム用の記述言語である AIML について、その記法の利点と問題点について述べた上で、隣接ペアの概念に基づいて独自に拡張し、隣接ペアの拡張形の情報を含むようにした会話シナリオ記述言語 AIML-ap を定義する。“ap” は隣接ペアの英語表記“Adjacency Pairs”の頭文字である。以下にその詳細を示す。

3.1 AIML

AIML は、A.L.I.C.E と呼ばれるチャットシステムの知識記述言語として、プログラムのサイズを小さくし、会話のためのデータを分離するために、xml を拡張する形で開発された [44]。A.L.I.C.E はもっとも人間に近い会話を行なった人工知能を表彰するローブナー賞を 3 度受賞しており、その記述言語である AIML は、それぞれ独自に拡張が行われて、現在でも様々なシステムで利用されている [45, 46]。AIML は図 3.1 のように、<pattern> タグで囲まれたシステムへの入力パターンと、<template> タグで囲まれたその入力に対応した出力のペアを持つ <category> タグを基本の単位として応答のための知識データベースを記述する。入力パターンにはワイルドカードが使用可能となっており、入力の揺れやユーザからのオープンな情報の入力を受け取ることが可能となっている。例えば図 3.1 の例をロードした AIML の処理系は、ユーザが「よく散歩に行くんだ」というテキストを入力した場合、AIML の処理系は「いつも行く場所とかあるんですか?」という出力を返す。

3.1.1 AIML の表現力

AIML では、その AIML ファイルがロードされている間は、原則全ての <category> タグが応答候補として利用される。しかし、これだけでは会話の進行に合わせた柔軟な会話内容の変更が行えない。そこで、AIML では以下のタグを定義することでその問題を軽減している。

```

<aiml>
  <category>
    <pattern>よく散歩に行くんだ</pattern>
    <template>いつも行く場所とかあるんですか？</template>
  </category>
  <category>
    <pattern>ない</pattern>
    <that>いつも行く場所とかあるんですか？</that>
    <template>そうなんですか</template>
  </category>
</aiml>

```

図 3.1: AIML の例

<set> タグ, <get> タグ

<set> タグと <get> タグは <template> タグの中で使用可能で, AIML 内で利用可能な変数へアクセスするためのタグである. <set> タグは, タグで囲った文字列を name プロパティの名前の変数に代入し, <get> タグは自分自身を name プロパティで指定した変数の中身に置換する. <set> タグと <get> タグはそのタグを含む <template> タグが返答で使用される直前に実行される. なお, <set> タグは実行後, 自身を代入した文字列に置き換えるため, 代入した内容を出力に含めたくない場合には, 囲った部分を出力に含めなくする <think> タグを利用する.

<topic> タグ

<topic> タグは, 複数の <category> タグを囲うタグで, <set> タグで代入できる変数空間の中の topic という名前の変数の値が <topic> タグの name プロパティで指定された文字列と一致した場合のみ, <topic> タグで囲まれた <category> 群を入出力の候補データとして使用する. なお, 同じ内容の <pattern> を持つ <category> タグが存在する場合, <topic> タグ内の Category が優先される.

<that> タグ

<that> タグは, <category> タグの直下で使用可能で, <that> タグを持つ <category> タグは直前の発話が <that> タグで囲まれたパターンに一致する場合のみ, その <category> タグは入出力の候補データとして使用される. なお, 同じ内容の <pattern> タグを持つ <category> タグが存在する場合, <that> タグを持つ Category が優先される. また <that> を含む <category> タグは <topic> 内の Category よりも優先される.

<condition> タグ

<condition> タグは, <template> タグの中で使用可能で, name プロパティで名前を指定した変数の値に応じて, <condition> タグをおきかえる文字列を決定する.

<star> タグ

<star> タグは <pattern> タグ内でワイルドカードを使用した場合に, index プロパティで指定したワイルドカードの中身を取得できる. <pattern> タグ内のワイルドカードは文頭側から順に番号が振られ最初のワイルドカードの番号は 1 である.

```

<aiml>
  <category name="main">
    <template>いつも行く場所とかあるんですか?</template>
    <pattern>ない</pattern>
    <template>そうなんですか</template>
    <next type="NONE">continue</next>
  </category>
</aiml>

```

図 3.2: AIML-ap の例

3.1.2 AIML の問題点

AIML は `<that>` タグなどを利用して柔軟な会話シナリオの表現を実現するが、以下のような問題も存在している。

1. 会話タスクのための知識データのフォーマットが「ユーザからの質問—システムの応答」という形式に限定されている
2. 挿入拡張のような会話の進行を記述しづらい。

1 つめは、AIML の成立過程に起因するもので、AIML が入力を受け取って出力を返すというチャットボットのための言語であるため、隣接ペアの概念でいえば、ユーザの入力が第 1 部分でシステムの出力が第 2 部分となるようなパターンしか `<category>` タグに記述できないという問題である。2 つめは、`<that>` タグが直前のシステムの出力を参照するため、隣接ペアにおける挿入拡張のような、システムからの質問に回答する前にユーザが一度質問を挟むような場合でも前後関係を維持するためには、ユーザからの質問の回答の `<that>` タグをもつ `<category>` タグを追加しなければならないという問題がある。

3.2 AIML-ap

AIML のデータ形式ではユーザの入力を先行発話、返答の出力を後続発話とした隣接ペアしか表現することができない。そこで、図 3.2 のように AIML を拡張することで隣接ペアの基本形及びその拡張を表現できるようにした。主な変更点は 3 点ある。まず `category` タグの中には最大で 3 つまで `<template>` タグと `<pattern>` タグを挿入でき、その順番がそのまま隣接ペアにおける第 1 部分、第 2 部分、第 3 部分を表す。次に隣接ペアの拡張を表現するため、`<category>` タグに `name` プロパティを追加した。また、`<next>` タグは後続する隣接ペアの `name` プロパティをマークアップし、後続する隣接ペアとの関係性を示す `type` プロパティを持つ。`<next>` タグの `type` プロパティは表 3.1 のように 5 種類用意した。したがって、`<next>` タグの `type` プロパティから隣接ペアの拡張形の範囲を知ることができる。

AIML-ap を用いた会話タスクはそれぞれ、変数 `CategoryName`、登録された `<category>` タグの集合、`<set>` タグや `<get>` タグなどで使用するローカル変数、最後に参照した `<next>` タグを内部変数として保持することで、システムの出力を決定することができる。以下に AIML-ap のデータセットを元に出力を決定する手順を示す。

1. ユーザから入力を受け取ると、変数 `CategoryName` の中身と同一の `name` プロパティを持つ `<category>` タグのうち、直近の入出力の履歴と第 1 部分ないしは第 1 部分と第 2 部分が一致する `<category>` タグを選択し、一致した部分の次の `<template>` タグを出力する

2. 相手からの入力、ないし `<template>` の出力を行なった段階でその `<category>` タグ内の全ての `<pattern>` タグおよび `<template>` タグが使われた場合、`<next>` の type プロパティが
 - (a) INSERT の場合は最後の出力を履歴から削除する
 - (b) INSERT でない場合は、変数 `CategoryName` の中身を `<next>` タグがマークアップする文字列に更新する．また、直近の入出力の履歴をリセットする
3. (2) で category name を更新した場合、変数 `CategoryName` の中身と同一の name プロパティを持つ `<category>` タグのうち、`<template>` タグで始まるものがあれば先頭の `<template>` タグを出力する

以上の手順により、`<pattern>` タグだけが異なるような `<category>` タグを複数用意した場合でも、正しく出力を決定することができる．

表 3.1: Next タグの Type プロパティの種類

| Type | 関係性 |
|--------|----------|
| PRE | 前方拡張 |
| POST | 後方拡張 |
| INSERT | 挿入拡張 |
| END | 会話タスクの終了 |
| NONE | 前後で関係なし |

3.2.1 AIML-ap の記述力

隣接ペアの拡張形の範囲が記述されている以外の AIML-ap の利点として、図 3.2 のようにロボット側が第 1 部分の発話を行うものや、第 3 部分を持つ隣接ペアを 1 つの `<category>` タグで表すことができることが挙げられる．AIML で記述しようとした場合、図 3.1 では 2 つに分割されてしまっている上、前後関係を指定するためには `<that>` タグを用いて直前のシステムの出力を指定しなければならず、本来の隣接ペアの関係が分かりづらくなってしまっている上に、記述には 9 行必要になる．しかし、AIML-ap で記述すれば 6 行で記述することができる．

また、隣接ペアの挿入拡張のような会話進行を AIML と AIML-ap で記述する場合、図 3.3 のようになる．例は、次の日に遊びに行く約束を取り付けるための質問を行なっている場面で、次の日の予定を尋ねる質問に対して、具体的には何時頃からの予定を訪ねているのかを質問し返している．AIML の場合では挿入拡張を追加する場合、それぞれの返答パターンにそれぞれ挿入拡張に対応した `<that>` タグを持つ入力を用意する必要があり、実用的ではない．しかし、AIML-ap では挿入拡張は会話のメインの流れとは独立して追加することができるので、`<category>` タグの数が無駄に増えることはない．図 3.3 の例では、AIML では 19 行、AIML-ap では 15 行で書くことができおり、もっともシンプルな場合でも AIML-ap の方が行数が削減されていることがわかる．

```

<aiml>
  <category>
    <pattern>お話ししよう</pattern>
    <template>明日は暇ですか？</template>
  </category>

  <category>
    <pattern>何時</pattern>
    <that>明日は暇ですか？</that>
    <template>12時以降です</template>
  </category>

  <category>
    <pattern>暇です</pattern>
    <that>明日は暇ですか？</that>
    <template>遊びに行きませんか？</template>
  </category>

  <category>
    <pattern>暇です</pattern>
    <that>12時以降です</that>
    <template>遊びに行きませんか？</template>
  </category>
</aiml>

```

(a) AIMLを用いた場合

```

<aiml>
  <category name="ask_schedule">
    <template>明日は暇ですか？</template>
    <pattern>暇です</pattern>
    <next type="PRE">invite_leisure</next>
  </category>

  <category name="ask_schedule">
    <pattern>何時</pattern>
    <template>12時以降です</template>
    <next type="INSERT" />
  </category>

  <category name="invite_leisure">
    <pattern>暇です</pattern>
    <that>12時以降です</that>
    <template>遊びに行きませんか？</template>
  </category>
</aiml>

```

(b) AIML-apを用いた場合

図 3.3: 挿入拡張の会話シナリオを追加した場合の AIML と AIML-ap の記述例

3.3 まとめ

本章ではチャットシステムの知識記述言語であった AIML を、隣接ペアの構造に基づく形に拡張し、会話タスクに含まれる隣接ペアの拡張形の情報を記述可能とした AIML-ap を定義した。AIML-ap は AIML と比べて、より簡潔に会話タスクのフローを記述することが可能となり、記述力の面も向上したものとなっている。

第 4 章

優先順位づけ因子を考慮した会話タスクスケジューリング手法 CACTS

4.1 優先順位づけ因子の候補

2 章では, Schegloff の考察 [37] に基づいて会話タスクの優先順位づけ因子の候補として以下をあげた.

1. 聞き手から見た割り込み者との人間関係
2. 会話の長さ
3. 割り込み時の会話タスク参加者の感情
4. 会話の重要度

しかし, これらが実際にどれほど会話の優先順位づけに影響しているかは明らかではなかった. そこで本章では, まず予備実験を行って会話タスクの優先順位に影響を因子について議論し, 後半では, 3 章で明らかにした隣接ペアの拡張形と会話タスクの切り替えタイミングの関係性と優先順位づけ因子を考慮した会話タスクスケジューリング手法の提案を行う.

4.2 各因子の影響力の分析

4.2.1 実験趣旨

会話外からの割り込みにおける優先順位づけ因子の候補が本当に優先順位づけに影響を与えているのか, また, 影響力があるならばそれはどのくらいかという点を明らかにしたい. そこで, 被験者に様々な割り込みシチュエーションを WEB ブラウザ上で提示し, 割り込まれた会話タスクと割り込みによって新たに生成された会話タスクのどちらを優先すべきかを判定してもらい, その傾向から因子候補の優先順位づけへの影響力を明らかにする.

4.2.2 因子の尺度と実験パターン

本研究で対象とした各因子候補はどれも数値化が困難なものであり, 連続尺度ではなく順序尺度や名義尺度で表現するのが望ましいと考えられる. そこで, 各因子候補について暫定的に表 4.1 のような名義尺度を定めた. 人間関係の因子以外は会話タスクごとに水準が設定される. 重要度については明確な達成すべき目的の有

表 4.1: 各因子候補の暫定尺度

| 因子 | 水準 |
|-------------------------------|------------------------------|
| 順序尺度 | 優先度低い ← - - - - → 優先度高い |
| 聞き手から見た割り込み者との人間関係 (Relation) | 0:下位, 1:同格, 2:上位 |
| 会話の長さ (Length) | 0:長い, 1:短い |
| 会話の重要度 (Importance) | 0:雑談, 1:緊急で無い用事, 2:緊急の用事 |
| 名義尺度 | |
| 割り込み時の会話タスク参加者の感情 (Emotion) | Joy, Neutral, Sadness, Anger |

無で雑談か用事かで水準を分け、さらに用事の中でも緊急性があるものかどうかで水準を分けている。ただし、人命に関わるタスクについて最優先に処理することは明らかだと考えられるので、そのようなタスクは緊急性のあるタスクからは除外したものについてのみ扱う。

さらに、実験ではこれら全ての組み合わせの会話の割り込みをすると検証すべきパターンが膨大になり、判断の条件が複雑化することで被験者の判断のブレが大きくなってしまいう危険性がある。そこで、3つの制限を加えて検証パターンの削減を行った。まず1つ目として、割り込む側・割り込まれる側の会話タスクで異なる水準を持つ(人間関係の場合は同格でない場合)因子は1つないし2つまでとし、水準が異なっている因子が2つある場合はそれぞれの会話タスクに優位と考えられる因子が存在するようにした。これは、2つの尺度間でどちらの因子が優先されるかを明らかにすれば、全ての因子候補間の影響力の順序を明らかにすることができるので2尺度間の比較で十分と考えられるためである。検証パターンの2つ目の条件として、それぞれの会話タスクの因子候補を揃える場合の水準はそれぞれ表 4.2 とした。最後の条件として、「雑談-緊急の用事」の組み合わせは除外した。これは、明確な達成すべき目的を持つか否か、は大きな差であると考えられるので「雑談-緊急でない用事」のみの比較で十分だと考えられるためである。結果、以下の条件を満たす会話タスクの組は 39 通りあり、それぞれ割り込み側と割り込まれる側を入れ替えた場合があるため、検証を行った割り込みのシチュエーションは 78 通りとなった。

表 4.2: 因子の水準を揃える場合の値

| 聞き手から見た 割り込み者との人間関係 | 会話の長さ | 会話の重要度 | 割り込み時の 会話タスク参加者の感情 |
|------------------------|-------|---------|-----------------------|
| 同格 | 短い | Neutral | 緊急度の低い用事 |

4.2.3 実験方法

実験のために図 4.1 のような web ブラウザ上で動作する UI を用意した。まず図 4.1 上が表示される。これは上から、話している人の地位、表情、およびその人の会話内容のわかる呼びかけの言葉が並べられており、一つの会話タスクの状況を表すセットとなっている。それぞれは以下の根拠に基づき用意した。

表情 尺度の上の感情を正しく被験者に伝えるため、Fischer らによって作成された表情セット ADFES[47] を用いた。ADFES は表情を動画で撮影したデータセットで、同時に「動画中のモデルの表情の頂点 (最も顔の筋肉が緊張している時) を切り出した^{*1}」画像も含んでいる。本実験では、会話内容について割り込みの表現に呼びかけ形式を用いたために喋らない動画では違和感が出てしまうことを考慮し、静止画の画像を用いた。実験には NorthEuropean の男性のうち M04,M06,M06 の id のモデルを用いた。

人間関係 表情の写真には人間関係に関する情報はないため、違和感のない設定として「上司」「部下」のラベルを用いた。またそれぞれのモデルに対して、M04 には上司、M06・M08 には部下のラベルを割り振った。

呼びかけ文 呼びかけ文は、筆者が用意した文章のうち、その会話の重要度と長さの因子について筆者を含めた大学生 3 名 (20 歳男性・20 歳男性・21 歳女性) が表 4.1 の尺度で分類し、全員の意見が一致した文章のみを用いた。また、選ばれた文章に対して表情の表す感情に合わせて言葉遣いに変化をつけたものを用意した。



図 4.1: 割り込みシチュエーション再現 UI

さらに、図 4.1 上が表示されてから 3 秒後に、右からもう一つの会話タスクの状況を表すセットが右からスライドして現れて図 4.1 下のような画面になる。このアニメーションが割り込みを表現している。

被験者はどちらの会話タスクに優先権があるか左右の矢印キーで選ぶ。ただし、上矢印キー、もしくはアニメーションの初めから 10 秒経過した場合は、選ぶのが困難であったとみなした。

^{*1} <http://aice.uva.nl/research-tools/adfes-stimulus-set/stimulus-set-contents/stimulus-set-contents.html> の ”freezing the film at the apex (point of most intense muscular contraction) of the model’s expression” の日本語訳

4.2.4 結果と考察

被験者は7名でいずれも大学生，男性が6名，女性が1名が参加した．平均年齢は20.1歳であった．また，全員から78のシチュエーション全ての回答を得た．回答のうち，23件はどちらを優先すべきかわからないという回答，あるいは時間内に回答できなかった．

わからない，あるいは時間内に回答できなかったできなかったとしたものを除いたデータ523件のうち，割り込み側を優先すると回答した割合は0.25だった．実験では，それぞれの会話タスクの因子の値は同じ組み合わせで，割り込む側と割り込まれる側かを入れ替えたものを両方実験しているので，立場による影響がないと仮定すると，どちらか一方で割り込む側を優先すると答えることになり，全データのうち割り込む側を優先すると回答する割合は50%に近似すると考えられる．割り込み側を優先すると回答する割合が0.5であるという仮説について，二乗検定をおこなった結果，二乗値は130でp値は0.0001以下となった．従って，割り込む側を優先すると回答する割合は明らかに50%ではないことから，割り込みの場面においては，割り込む側よりも割り込まれる側の方が優先されやすいことがわかった．

判定不能とした回答を除いたデータ523件を用いて多項ロジスティック回帰分析を行なった．本分析では，各因子の尺度は全て連続尺度とみなすことにする．割り込み側を優先する場合を1，しない場合を0として，これを目的変数とした．説明変数は，二つの会話タスク間の因子の差を表す変数として表4.3に示した人間関係(x_1)，長さの差(x_2)，目的の有無の差(x_3)，緊急度の差(x_4)，怒りの差(x_5)，悲しみの差(x_6)，喜びの差(x_7)の7つの説明変数を用いた．それぞれ，割り込まれた側を基準として，-1, 0, 1のいずれかの値をとる．

表 4.3: 説明変数と計算方法

| 変数名 | 計算式 |
|-----------------|---|
| x_1 (人間関係) | $Relation - 1$ |
| x_2 (会話長の差) | $Length_{\text{割り込んだ側}} - Length_{\text{割り込まれた側}}$ |
| x_3 (目的の有無の差) | (if $Importance_{\text{割り込んだ側}} = 0$ then 0 else 1) - (if $Importance_{\text{割り込まれた側}} = 0$ then 0 else 1) |
| x_4 (緊急度の差) | (if $Importance_{\text{割り込んだ側}} = 2$ then 1 else 0) - (if $Importance_{\text{割り込まれた側}} = 2$ then 1 else 0) |
| x_5 (怒りの差) | (if $Emotion_{\text{割り込んだ側}} = \text{Anger}$ then 1 else 0) - (if $Emotion_{\text{割り込まれた側}} = \text{Anger}$ then 1 else 0) |
| x_6 (悲しみ度の差) | (if $Emotion_{\text{割り込んだ側}} = \text{Sad}$ then 1 else 0) - (if $Emotion_{\text{割り込まれた側}} = \text{Sad}$ then 1 else 0) |
| x_7 (喜び度の差) | (if $Emotion_{\text{割り込んだ側}} = \text{Joy}$ then 1 else 0) - (if $Emotion_{\text{割り込まれた側}} = \text{Joy}$ then 1 else 0) |

最尤法を用いて多項ロジスティック回帰分析をおこなったところ，表4.4の結果を得た．本研究では，説明変数の説明力と影響力を明らかにするために使用するのでモデルの適合度については検討しない．p値が0.05以下の因子は x_1 (人間関係)， x_2 (長さの差)， x_3 (目的の有無の差)， x_5 (怒りの差)の4つで，影響力の強さは x_3 (目的の有無の差) > x_2 (長さの差) > x_1 (人間関係) > x_5 (怒りの差)の順となっている．この4つ以外の説明変数はp > 0.20となっており，説明力は低いと言える．

表 4.5 は、実験で検証したシチュエーションと、そのシチュエーションで割り込み側がどれだけ優先されたかの割合を示した。シチュエーションは割り込み側が優先された割合が高い順に並べられており、考察の際の参照を楽にするためにシチュエーションを表の順で id を振ってある。例えば、id1 の例では、目的の有無が 1、悲しみの差が-1、喜びの差が 1 となっていることから、割り込まれた側は目的がなく雑談をしており、悲しそうである。一方割り込み側は目的のある用事達成しようと割り込んできており、喜んだ表情をしている。また、二人は上下の関係にはなく、どちらも短めの会話タスクであるとわかる。

表 4.4: ロジスティック回帰分析の結果

(* : $p < 0.05$, *** : $p < 0.001$)

| 項 | 偏回帰係数 | 標準誤差 | 尤度比カイ 2 乗 | p 値 ($Prob > ChiSq$) |
|-----------------|---------|-------|-----------|------------------------|
| 切片 | -1.37 | 0.125 | 163.20 | < .0001*** |
| x_1 (人間関係) | 0.928 | 0.273 | 12.01 | 0.0005*** |
| x_2 (長さの差) | 1.23 | 0.278 | 1.640 | < .0001*** |
| x_3 (目的の有無の差) | 2.24 | 0.316 | 61.24 | < .0001*** |
| x_4 (緊急度の差) | 0.351 | 0.275 | 1.640 | 0.2004 |
| x_5 (怒りの差) | 0.634 | 0.273 | 5.465 | 0.0194* |
| x_6 (悲しみ度の差) | -0.287 | 0.235 | 1.507 | 0.2196 |
| x_7 (喜び度の差) | -0.0900 | 0.233 | 0.1492 | 0.6993 |

表 4.5: 各シチュエーションでの割り込み側が優先された割合

| ID | 人間関係 | 会話長の差 | 目的の有無 | 会話長の差 | 怒りの差 | 悲しみの差 | 喜びの差 | 割り込み側が優先された割合 |
|----|------|-------|-------|-------|------|-------|------|---------------|
| 1 | 0 | 0 | 1 | 0 | 0 | -1 | 1 | 6人 / 7人 |
| 2 | 0 | 0 | 1 | 0 | 0 | -1 | 0 | 6人 / 7人 |
| 3 | 0 | 0 | 1 | 0 | -1 | 0 | 1 | 5人 / 7人 |
| 4 | 0 | 0 | 1 | 0 | -1 | 0 | 0 | 5人 / 7人 |
| 5 | -1 | 0 | 1 | 0 | 0 | 0 | 0 | 5人 / 7人 |
| 6 | 0 | 1 | 0 | 0 | 0 | -1 | 0 | 5人 / 7人 |
| 7 | 0 | 1 | 0 | 0 | -1 | 0 | 1 | 4人 / 7人 |
| 8 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 4人 / 7人 |
| 9 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 4人 / 7人 |
| 10 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 4人 / 7人 |
| 11 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 4人 / 7人 |
| 12 | 0 | 1 | 0 | -1 | 0 | 0 | 0 | 3人 / 6人 |
| 13 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 3人 / 6人 |
| 14 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3人 / 7人 |
| 15 | 0 | 0 | 1 | 0 | -1 | 1 | 0 | 3人 / 7人 |
| 16 | 0 | 0 | 0 | -1 | 1 | 0 | 0 | 3人 / 7人 |
| 17 | 0 | 0 | -1 | 0 | 1 | 0 | -1 | 3人 / 7人 |
| 18 | 0 | 0 | 0 | 1 | 0 | -1 | 1 | 3人 / 7人 |
| 19 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3人 / 7人 |
| 20 | 0 | 0 | 0 | 1 | 0 | -1 | 0 | 3人 / 7人 |
| 21 | 1 | 0 | 0 | 0 | 0 | -1 | 0 | 3人 / 7人 |
| 22 | 0 | 1 | 0 | 0 | 0 | -1 | 1 | 2人 / 6人 |
| 23 | 0 | 1 | 0 | 0 | -1 | 0 | 0 | 2人 / 6人 |
| 24 | 0 | 0 | 0 | -1 | 1 | 0 | -1 | 2人 / 6人 |
| 25 | 0 | 0 | -1 | 0 | 1 | 0 | 0 | 2人 / 6人 |
| 26 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 2人 / 6人 |
| 27 | 0 | 1 | 0 | 0 | -1 | 1 | 0 | 2人 / 7人 |
| 28 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 2人 / 7人 |
| 29 | 0 | 0 | 0 | 0 | 1 | 0 | -1 | 2人 / 7人 |
| 30 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2人 / 7人 |
| 31 | 1 | 0 | 0 | 0 | -1 | 0 | 1 | 2人 / 7人 |
| 32 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 2人 / 7人 |
| 33 | 1 | 0 | 0 | 0 | 0 | -1 | 1 | 2人 / 7人 |
| 34 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 2人 / 7人 |
| 35 | 0 | 0 | 0 | -1 | 1 | -1 | 0 | 1人 / 6人 |
| 36 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1人 / 6人 |
| 37 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 1人 / 6人 |
| 38 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 1人 / 6人 |
| 39 | 0 | 0 | -1 | 0 | 1 | -1 | 0 | 1人 / 6人 |
| 40 | 0 | 0 | 0 | 1 | -1 | 0 | 1 | 1人 / 6人 |
| 41 | 1 | 0 | 0 | 0 | -1 | 1 | 0 | 1人 / 6人 |
| 42 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1人 / 7人 |
| 43 | 0 | 0 | 0 | -1 | 0 | 0 | -1 | 1人 / 7人 |

表 4.5: 各シチュエーションでの割り込み側が優先された割合 (続き)

| ID | 人間関係 | 会話長の差 | 目的の有無 | 会話長の差 | 怒りの差 | 悲しみの差 | 喜びの差 | 割り込み側が優先された割合 |
|----|------|-------|-------|-------|------|-------|------|---------------|
| 44 | 0 | -1 | 0 | 1 | 0 | 0 | 0 | 1人 / 7人 |
| 45 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 1人 / 7人 |
| 46 | 0 | -1 | 0 | 0 | 1 | 0 | -1 | 1人 / 7人 |
| 47 | 0 | -1 | 0 | 0 | 1 | 0 | 0 | 1人 / 7人 |
| 48 | 0 | -1 | 0 | 0 | 0 | 0 | -1 | 1人 / 7人 |
| 49 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 1人 / 7人 |
| 50 | 1 | 0 | 0 | 0 | -1 | 0 | 0 | 1人 / 7人 |
| 51 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1人 / 7人 |
| 52 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1人 / 7人 |
| 53 | -1 | 0 | 0 | 0 | 0 | 1 | -1 | 1人 / 7人 |
| 54 | -1 | 0 | 0 | 0 | 0 | 1 | 0 | 1人 / 7人 |
| 55 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 1人 / 7人 |
| 56 | 0 | 0 | 0 | -1 | 0 | 1 | -1 | 1人 / 7人 |
| 57 | 0 | 0 | 0 | 1 | -1 | 1 | 0 | 1人 / 7人 |
| 58 | -1 | 0 | 0 | 0 | 1 | -1 | 0 | 0人 / 6人 |
| 59 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0人 / 6人 |
| 60 | 0 | -1 | 0 | 0 | 0 | 1 | -1 | 0人 / 6人 |
| 61 | 0 | 0 | 0 | 0 | -1 | 1 | 0 | 0人 / 6人 |
| 62 | -1 | 0 | 0 | 0 | 1 | 0 | -1 | 0人 / 6人 |
| 63 | -1 | 0 | 0 | 0 | 1 | 0 | 0 | 0人 / 6人 |
| 64 | 0 | -1 | 0 | 0 | 1 | -1 | 0 | 0人 / 7人 |
| 65 | 0 | 0 | 0 | 0 | -1 | 0 | 1 | 0人 / 7人 |
| 66 | 0 | 0 | 0 | 0 | 0 | -1 | 1 | 0人 / 7人 |
| 67 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0人 / 7人 |
| 68 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0人 / 7人 |
| 69 | -1 | 0 | 0 | 0 | 0 | 0 | -1 | 0人 / 7人 |
| 70 | -1 | 0 | 0 | 1 | 0 | 0 | 0 | 0人 / 6人 |
| 71 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0人 / 7人 |
| 72 | 0 | -1 | 0 | 0 | 0 | 1 | 0 | 0人 / 7人 |
| 73 | 0 | 0 | -1 | 0 | 0 | 0 | -1 | 0人 / 7人 |
| 74 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0人 / 6人 |
| 75 | 1 | 0 | -1 | 0 | 0 | 0 | 0 | 0人 / 6人 |
| 76 | 0 | 0 | 0 | -1 | 0 | 1 | 0 | 0人 / 7人 |
| 77 | 0 | 0 | -1 | 0 | 0 | 1 | -1 | 0人 / 7人 |
| 78 | 0 | 0 | -1 | 0 | 0 | 1 | 0 | 0人 / 7人 |

いくつかのシチュエーションの割り込み側が優先された割合は、「目的の有無の差」>「長さの差」>「人間関係」>「怒りの差」という偏回帰係数の大小関係では説明できないものがある。そこで以下では、影響力がないと考えられる因子や多項ロジスティック回帰分析の結果と一致しないような現象について、その理由を考察する。

喜び度の差・悲しみ度の差

これらの説明力が低かった理由として、怒りの感情を表出している場合、割り込まれた本人は怒りの原因が自身に起因する可能性から解決を試みる必要性を感じるのに対し、喜びや悲しみの場合はタスクの達成の必要性を訴える要素がないため、そのような感情を表出している人を優先しようという動機に結びつかなかったと考えられる。従って、喜び度の差・悲しみ度の差については因子からは除外することとする。

緊急度の差

緊急度の差の説明力が低かった理由として、命に関わるような緊急性のあるものについては除外したため、命に関わらない用事の中では緊急度に大きな差を感じられなかったためだと考えられる。また、会話タスクはショッピングモールを想定したものを用意したため、「トイレの場所を知る」といったタスクについては人に尋ねる以外にも、案内地図といったメディアを使うことが考えられることから優先度が下がっていた可能性が考えられる。従って、緊急度の差については因子からは除外することとする。

「怒っている」かつ「雑談」の場合

割り込み側が「怒っている」かつ「雑談」の場合にも、割り込まれる側が「用事」の場合でも7人中3人と半数近い人数が割り込み側が優先したシチュエーション (ID:17) が存在している。多項ロジスティック回帰分析の結果では「目的の有無の差」>「怒りの差」であり、偏回帰係数の絶対値も大きく離れているため、直感に反する結果となっている。原因として、「怒っている」状態の人が会話を切り出したために、平常時ならば「雑談」であった内容がクレームという「用事」と捉えられてしまった可能性が考えられる。クレームは話をする事自体が目的化されてしまった結果、「用事」かつ「怒っている」と認識されるため、「用事」かつ「怒っていない」割り込まれ側よりも優先されやすいことが説明できる。ただし一方で、影響力があると考えられる因子について割り込む側と割り込まれる側で入れ替わっているシチュエーションにおいて、7人中3人が割り込む側を優先しており差がないものがある (ID:15)。このように、クレームとその他の用事ではどちらを優先すべきかの判断は難しさがあることは留意する必要がある。なお、「怒っている」場合の「雑談」としては内容が愚痴である場合も考えられるため、怒っているならば一概に「用事」であるとは言えないと考えられる。

「会話の長さ」と「人間関係」の影響力

「会話の長さ」と「人間関係」が対立しているシチュエーション (ID:37,67) において、割り込み側にとって「人間関係」が優位で「会話の長さ」が不利な場合 (ID:37) の方が割り込み側が逆の場合 (ID:67) より割り込み側が優先された割合は高くなっている。しかし、この二者の間の Fisher の正確確率は $p=0.462$ で有意な結果とは言えない。従って、多項ロジスティック回帰分析の結果を覆すような結果ではないと考えられる。

4.2.5 まとめ

予備実験の結果、各因子の影響力の大小関係は「目的の有無の差」>「長さの差」>「人間関係」>「怒りの差」であると推定され、この4つの因子以外は影響力が小さく、無視することが可能であると考えられる。また、クレームという会話タスクは、会話することそのものが自己目的した結果、会話の終着点が見えないような場合でも目的がある会話タスクと同等となり得るので、目的有りとして扱うべきではないかという知見が得られた。

4.3 会話タスクスケジューリング手法 CACTS

CACTS は、会話タスクを実行中のロボットが会話外からの割り込みを受けるなどして複数の会話タスクの実行を要求された際に、適切なタイミングでの会話タスクの移行を行うための会話タスクスケジューリング手法である。CACTS という名前は、Context Aware Conversation Task Scheduling の頭文字をとったもので、“Context” という単語は、会話タスクのスケジューリングに用いている「会話の文脈」と「割り込みの状況」のダブルミーニングとなっている。「会話の文脈」とは、ここでは会話タスク中に現れる隣接ペアの拡張形のことを表している。3章で議論したように、隣接ペアの拡張形は会話タスクの切り替えタイミングに制約を与えており、会話タスクのスケジューリングを行う上で重要な情報である。一方「割り込みの状況」とは、割り込んだ会話タスクの優先度を表しており、CACTS では「会話タスク間の目的の有無の差」、「会話タスク間の長さの差」、「話者同士の間関係」、「話者の怒りの感情の差」の4つの因子に基づいて割り込んだ会話タスクの優先度を決定する。そして、新たな会話タスクへの切り替えを優先度に応じて、割り込み内容を確認した後に実行するのか、隣接ペアの拡張形の完成を待つのか、あるいは、現在の会話タスクが終了するのを待つのかを決定する。また、優先度が同一の会話タスクが存在する場合には、会話タスクが生成された時間が早い方を優先して切り替える。

本節では、会話外からの割り込みによって生成される会話タスクのスケジューリングを行うための手法 CACTS の詳細について述べる。まず、CACTS の核となる会話タスクの切り替えの判断基準を示し、前節の実験結果との整合性について確認する。その後、CACTS の会話タスクのスケジューリングアルゴリズムを示す。

4.3.1 会話タスクの切り替えの判断基準

前節で議論したように、会話タスクの優先度には「会話タスク間の目的の有無の差」、「会話タスク間の長さの差」、「話者同士の間関係」、「話者の怒りの感情の差」の4つの因子が影響を与えていることがわかっている。本手法では、各因子を用いた線形な計算式によって優先度を決定する。

予備実験で調査したシチュエーションのうち、50%以上の割合で割り込み側が優先されたシチュエーションでは割り込まれた側から一定の理解を得られたとみなし、それらのシチュエーションが会話終了直後に切り替わるような優先度の計算式および閾値として、表4.6のようなパラメータを与えた。

表 4.6: CACTS における優先度の計算式と切り替えタイミングの閾値

| 優先度の差 | | $dP_{AB} = \text{目的の有無の差} * 9 + \text{会話長の差} * 5 + \text{人間関係} * 3 + \text{怒りの差} * 1$ |
|--------------------|---------------|---|
| 切替 タイミング の閾値 | 割り込み内容を確認した後 | $dP_{AB} > 3$ |
| | 隣接ペアの拡張形の完成直後 | $dP_{AB} > 0$ |
| | 会話タスクの終了後 | $dP_{AB} \leq 0$ |

表 4.6 を適用した場合、50%以上の割合で割り込み側が優先された 13 シチュエーションのうち 12 シチュエーションの切り替えタイミングを「割り込み内容を確認した後」と判定し、全 78 実験シチュエーションのうち 17 シチュエーションの切り替えタイミングを「割り込み内容を確認した後」と判定する。以下で判定が

一致しないシチュエーションについて整合性を検討する。

割り込み側が怒っている場合

割り込み側が「怒りの差」で優位でその他の影響力が認められた因子に差がないシチュエーション (ID:11) は、50% 以上の割合で割り込み側が優先されている。しかしながら、その他の「怒りの差」が優位でその他の影響力が認められた因子に差がないシチュエーションでは 50% 以上の割合で割り込み側が優先されたものはない。また、「怒りの差」より強い影響力を持つ「人間関係」が優位の場合でも、その他の影響力が認められた因子に差がないシチュエーションでも 50% 以上の割合で割り込み側が優先されたものはないことから、本シチュエーションは外れ値として扱うべきだと考えられる。

「怒りの差」と「目的の有無の差」

50% 以上の割合で割り込み側が優先されていないが「割り込み内容を確認した後」と判定されたシチュエーションの一つは、「怒りの差」で優位な「雑談」に対して「用事」が割り込むシチュエーション (ID:15) である。これは、割り込まれた側が、前小節で考察したようにクレームと認識された「怒っている」状態の「雑談」と「用事」との優先づけの判断の困難さに起因していると考えられる。クレームと判断できる会話タスクを「目的有り」とみなした場合、このシチュエーションでの会話タスクの切り替えタイミングは「会話タスク終了後」となる。これは、7人中3人が割り込み側を優先したことを考えると切り替えタイミングが遅いとも考えられるが、「怒っている」状態の「用事」に対して「用事」が割り込んだ場合 (ID:65) には全会一致で割り込まれた側を優先していることから、妥当であると考えられる。

「会話長の差」の影響力と閾値

残りの 50% 以上の割合で割り込み側が優先されていないが「割り込み内容を確認した後」と判定されたシチュエーション (ID:14,22,23,27) は全て「会話長の差」が割り込み側に優位である。影響力が認められた因子についてのみ注目した場合、ID:14,22 と ID:6, ID:23,27 と ID:7 のシチュエーションが一致する。本研究では 50% 以上の割合で割り込み側が優先されたという事実に着目し、判定の再現率を上げるためこのような不一致を許容した。

4.3.2 会話タスクスケジューリングの流れ

ここでは、会話タスク切り替え手法 CACTS のアルゴリズムについて説明する。まず、会話タスクの切り替えタイミングの判断基準のために会話タスクが保持すべきデータおよび、そのデータに関する関数をアルゴリズム 1 に示した。

Algorithm 1 スケジューリングに必要なデータおよび関数

Require: 会話タスクは 重要度 (0 or 1) , 長さ (0 or 1) , 話者の怒り (0 or 1) , 話者 , 生成された時刻 の 情報を持つ

```
1: function 人間関係 (話者 A, 話者 B)
2:   if 話者 A が話者 B より上位 then
3:     Return 1
4:   else if 話者 A と話者 B が同格 then
5:     Return 0
6:   else if 話者 A が話者 B より下位 then
7:     Return -1
8:   else
9:     Return 0
10:  end if
11: end function
```

本手法の会話タスクの切り替えタイミングそれぞれについて、「割り込み内容を確認した後」の会話タスクの切り替えアルゴリズムを Algorithm2 に、「初めて隣接ペアの拡張形が完成した後」の会話タスクの切り替えアルゴリズムを Algorithm3 に、「現在の会話タスクが終了した後」の会話タスクの切り替えアルゴリズムを Algorithm4 に示す。

Algorithm 2 割り込み内容を確認した後の会話タスク切り替えフロー

```
1: procedure 割り込み内容を確認した後の会話タスク切り替え
2:    $C =$  現在実行中の会話タスク
3:    $N =$  割り込みによって追加された会話タスク
4:    $N$  の優先度 =  $(N$  の重要度 -  $C$  の重要度) * 9 +  $(N$  の長さ -  $C$  の長さ) * 5 + 人間関係 ( $N$  の話者,  $C$  の話者) * 3 +  $(N$  の話者の怒り -  $C$  の話者の怒り)
5:   if  $N$  の優先度 > 3 then
6:     現在の会話タスクを中断
7:     中断している会話タスクのリスト +=  $C$ 
8:     現在実行中の会話タスク =  $N$ 
9:     中断している会話タスクのリスト -=  $N$ 
10:  else
11:    中断している会話タスクのリスト += 割り込みによって追加された会話タスク
12:  end if
13: end procedure
```

Algorithm 3 隣接ペアの拡張形が完成した後の会話タスク切り替えフロー

```
1: procedure 隣接ペアの拡張形が完成した後の会話タスク切り替え
2:    $C =$  現在実行中の会話タスク
3:    $N = \text{NULL}$ 
4:    $N$  の優先度  $= 0$ 
5:   for 中断している会話タスク in 中断している会話タスクのリスト do
6:      $S =$  中断している会話タスク
7:      $S$  の優先度  $= (S$  の重要度  $- C$  の重要度)  $* 9 + (S$  の長さ  $- C$  の長さ)  $* 5 +$  人間関係 ( $S$  の話者,  $C$  の話者)  $* 3 + (S$  の話者の怒り  $- C$  の話者の怒り)
8:     if  $S$  の優先度  $> N$  の優先度 then
9:        $N = S$ 
10:       $N$  の優先度  $= S$  の優先度
11:     else if  $S$  の優先度  $== N$  の優先度 then
12:       if  $S$  が生成された時刻  $< N$  の生成された時刻 then
13:          $N = S$ 
14:          $N$  の優先度  $= S$  の優先度
15:       end if
16:     end if
17:   end for
18:   if  $N$  が  $\text{NULL}$  じゃない then
19:     現在の会話タスクを中断
20:     中断している会話タスクのリスト  $+= C$ 
21:     現在実行中の会話タスク  $= N$ 
22:     中断している会話タスクのリスト  $-= N$ 
23:   end if
24: end procedure
```

Algorithm 4 現在の会話タスクが終了した後の会話タスク切り替えフロー

```
1: procedure 現在の会話タスクが終了した後の会話タスク切り替え
2:    $C =$  現在実行中の会話タスク
3:    $N = \text{NULL}$ 
4:    $N$  の優先度 = NULL
5:   for 中断している会話タスク in 中断している会話タスクのリスト do
6:      $S =$  中断している会話タスク
7:      $S$  の優先度 =  $(S$  の重要度 -  $C$  の重要度) * 9 +  $(S$  の長さ -  $C$  の長さ) * 5 + 人間関係 ( $S$  の話者,  $C$  の話者) * 3 +  $(S$  の話者の怒り -  $C$  の話者の怒り)
8:     if  $N$  が NULL then
9:        $N = S$ 
10:       $N$  の優先度 =  $S$  の優先度
11:     else if  $S$  の優先度 >  $N$  の優先度 then
12:        $N = S$ 
13:       $N$  の優先度 =  $S$  の優先度
14:     else if  $S$  の優先度 ==  $N$  の優先度 then
15:       if  $S$  が生成された時刻 <  $N$  の生成された時刻 then
16:          $N = S$ 
17:          $N$  の優先度 =  $S$  の優先度
18:       end if
19:     end if
20:   end for
21:   現在の会話タスクを終了
22:   現在実行中の会話タスク =  $N$ 
23:   中断している会話タスクのリスト -=  $N$ 
24: end procedure
```

4.4 まとめ

本章では、割り込みが生じた際に割り込んだ側と割り込まれた側のどちらを優先するかを被験者に選ばせたデータをもとに、まず会話タスクの優先順位に影響を因子について議論し、優先順位への影響力の高い順に「目的の有無の差」「長さの差」「人間関係」「怒りの差」の因子が影響していることがわかった。さらに、これらの因子を考慮して3章で議論した会話構造に基づくタイミングでの会話の切り替えを行う会話タスクスケジューリング手法 CACTS を定義した。

第 5 章

AIML-ap を用いた CACTS の実装と評価

本章では、会話タスクスケジューリング手法 CACTS の有効性について議論を行う。まず、CACTS を用いた場合のインタラクションについて評価を行うために、AIML-ap を用いて実装したロボットシステムについて述べる。次に、CACTS の客観的な立場での印象についての実験の詳細について述べる。さらに、被験者が割り込まれる立場になった時の CACTS を用いたロボットの印象について評価する実験について述べ、CACTS の有効性について議論する。

5.1 実験システムの概要

CACTS を用いたロボットとインタラクションを行う実験を行うため、これまでの議論を踏まえて、CACTS を用いた会話タスク実行ロボットシステムを Softbank Robotics 社の Pepper 上に実装した。システム構成図を図 5.1 に示す。本節では、まず本システムにおける会話タスクの内部表現について示し、その後システムの各部分の機能について説明する。

5.1.1 会話タスクの内部表現

会話タスクには「応答生成用のデータ」と「5 章で定義した CACTS が要求する情報」が含まれる。本システムの応答生成のためのデータは、4 章で定義した AIML-ap で記述される。また、会話タスクで使用するローカル変数も含まれる。また、「5 章で定義した CACTS が要求する情報のうち、「会話タスクの目的の有無」「会話タスクの長さ」の 2 つは AIML-ap スクリプトのメタデータとしてあらかじめ記述されている。また、「人物情報」「感情」「生成された時刻」は会話タスク生成時に付与される。

5.1.2 入出力インタフェース

本実装では、Pepper 上に実装されているミドルウェアである NAOqi と、web カメラを搭載したシングルボードコンピュータ raspberry pi 3 Model B を使用した。本システムでは、NAOqi の機能のうち次の音声認識機能、顔認識機能、顔追跡機能および音声合成機能を使用した。

音声認識機能 NAOqi の音声認識機能は、入力音声事前に登録した単語かどうかの確信度を計算する。本システムでは会話の進行に合わせて逐次会話タスク実行モジュールから渡される入力候補の単語を音声認識の候補単語として利用し、その結果を入力認識モジュールに送っている。また、入力音声の音源位

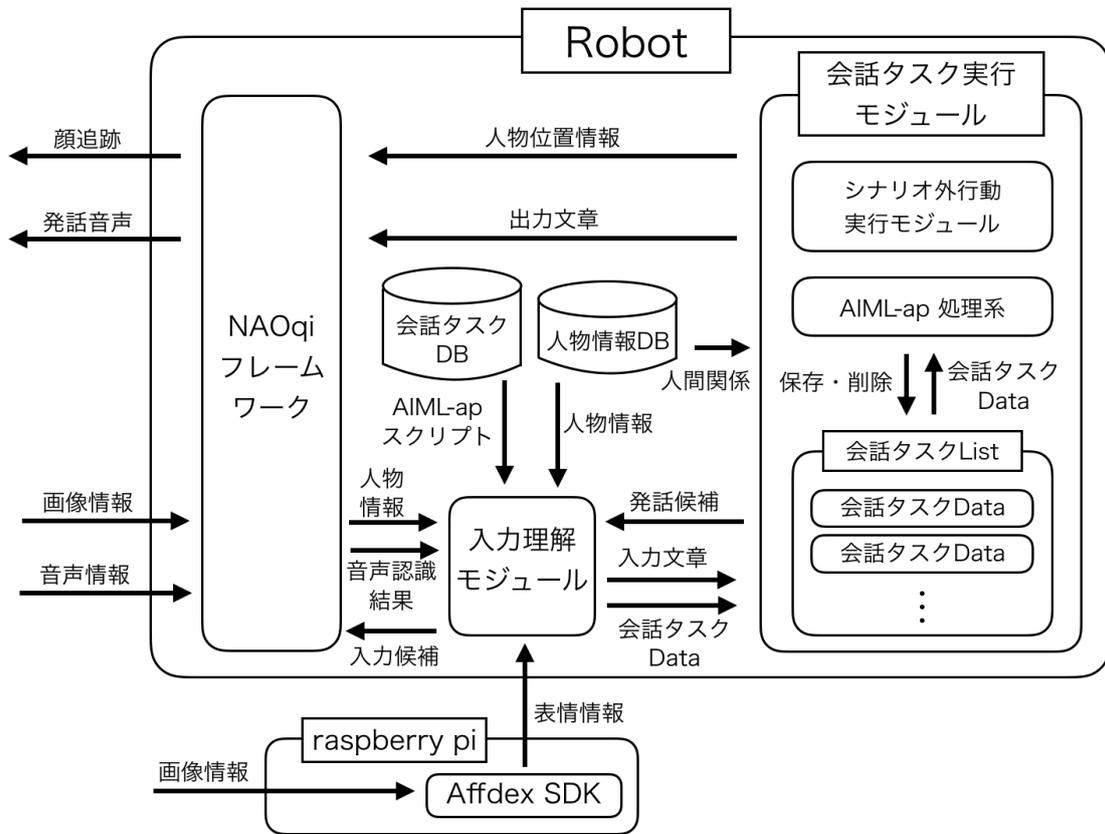


図 5.1: システム構成図

置の推定を行っており，正面から一定の角度外れた方向か否かを判定しその結果を入力認識モジュールに送っている．

顔認識機能 カメラ内の顔を発見，および事前に登録された顔と照合し，その結果を入力認識モジュールに送っている．

顔追跡機能 本システムでは，会話相手の方を安定して向くため顔追従機能を使用している．また，音声認識機能にて入力音声の方向が正面から一定の角度外れている場合や，会話相手を変更する場合には次に発話すべき相手に顔追跡の対象を変更する．

音声合成モジュール 会話タスク実行モジュールの出力結果を音声としてスピーカから出力する．

また，会話相手の表情情報をセンシングするために raspberry pi 3 Model B を使用した．raspberry pi 上で動作する表情認識エンジン Affdex SDK*1 は顔画像から表情の推定を行い，表情に含まれる「Joy」「Sad」「Anger」などの感情の強さを 0 から 100 の間の値として出力する．本システムでは Pepper の頭部に取り付けた web カメラの画像から得られた，「Anger」の推定結果を閾値で怒っているか否かの 2 値に変換して入力認識モジュールに送っている．

*1 <https://www.affectiva.com/product/emotion-sdk/>

5.1.3 入力処理モジュール

入力処理モジュールは次の機能を担当する。

1. 音声認識の候補単語の決定
2. 音声認識結果と会話の文脈に基づく入力結果および発話意図の推定
3. 会話タスクの生成

音声認識機能の拡張

まず、本システムが受け取る入力は以下の3種類が存在する。

現在の会話タスクにおける返答候補 現在遂行されている会話タスクの会話相手の返答候補は、会話の進行に応じて会話タスク実行モジュールによって逐次更新され、会話タスクの進行のために用いられる。

割り込みの開始および誤認識を検出する単語 割り込み前後の特徴的な単語として開始のトリガーとして「ねえねえ」および「すみません」を、また割り込み誤認識を示す単語として「なんでもない」を設定した。

要求された会話タスクを理解するための単語 割り込み発生後の会話タスクの要求を認識するために用いられる。単語候補は会話タスクデータベースの各会話タスクデータから取得する。

音声認識の精度を向上させるため、入力処理モジュールの会話の状態を「平常」「割り込み中」に分け、状態に応じて3種類のうち音声認識の候補として用いる単語を選択し、認識候補の単語を減らすようにした。次の状態は、現在の状態と入力文の種類に応じて決定される。さらに、NAOqiの仕様では、音声認識の結果は候補となっている単語それぞれの確信度のスコアとして返されるので、「単語の種類」と「会話の状態」、および「声する方向」に応じて確信度のスコアにボーナスのスコアを加算し、最終的なスコアが閾値を超えたもののうち最大のスコアを持つ単語を入力文と判定する。声の方向によるボーナスの設定によって、周辺のノイズによる誤認識を防ぐことができると考えられる。閾値を超えるスコアを持つ単語が存在しなかった場合は、音声認識失敗とみなし「えっ？」という聞き返しを行い、同じ会話の状態ですべての単語の選定から再実行する。会話の状態に応じた候補単語と確信度スコアのボーナスは表5.1のように設定した。

表 5.1: 会話の状態と音声認識の処理

| 会話の状態 | 音声認識の候補単語 | ボーナスの条件 | 次の会話の状態 |
|-------|----------------------|-----------------------|---------|
| 平常 | 現在の会話タスクにおける返答候補 | 正面からの音声ならばスコアにボーナスを加算 | 平常 |
| | ねえねえ/すみません | 側面からの音声ならばスコアにボーナスを加算 | 割り込み中 |
| 割り込み中 | なんでもない | 正面からの音声ならばスコアにボーナスを加算 | 平常 |
| | 要求された会話タスクを理解するための単語 | 正面からの音声ならばスコアにボーナスを加算 | 平常 |

判定した入力文は会話タスク実行モジュールに送る。

会話タスクの生成

入力文が要求された会話タスクを理解するための単語であった場合、その単語の属する会話タスクの AIML-ap データと raspberry pi 3 Model B からの表情情報、顔認識機能の結果を用いて人物情報データベースを検索して得られた人物情報を統合して会話タスクデータを生成し、会話タスク実行モジュールに送る。

5.1.4 会話タスク実行モジュール

会話タスク実行モジュールは、現在実行されている会話タスクの AIML-ap スクリプトに基づき AIML-ap の処理系が入力文の処理をおこない、NAOqi フレームワークを用いて音声として出力する。また、会話タスク実行モジュールには CACTS が実装されており、「会話タスクが新たに生成された後（「割り込み内容を確認した後」に相当するタイミング）」「割り込み内容を確認した後」「初めて隣接ペアの拡張形が完成した後」「現在の会話タスクが終了した後」には会話のスケジューリングが実行される。

割り込みによって会話タスクが会話タスクリストに追加された場合、会話タスク内に含まれる会話タスクスケジューリングのためのデータと人物情報データベースの人間関係情報を用いて、会話タスクのスケジューリングを行う。また、AIML-ap の処理系による出力以外の行動についてはシナリオ外行動実行モジュールを用いる。具体的には、以下のような補完を行う。

シナリオ外行動実行モジュール

1. 入力文の種類が「割り込みの開始および誤認識を検出する単語」だった場合、単語に対応する定型文を返す。
2. 会話中にスケジューリングされていた会話の切り替えが実行された場合、待ってもらった側や待ってもらった側に定型文の発話を行う。

5.1.5 システムの動作の流れ

ロボットが会話相手の発話を音声情報として受け取ると、NAOqi フレームワークを介して音声認識結果と音源定位情報を入力理解モジュールが受け取る。入力理解モジュールは、現在の会話の状態と音源定位情報を元に単語ごとの確信度の調整を行い、閾値を超える候補単語があった場合、入力の種類に応じて、入力を会話タスク実行モジュールに送るか、会話の状態を変化させるか、新しい会話タスクを生成して会話タスク実行モジュールに送るかを決定する。会話タスク実行モジュールは、入力として単語を受け取った場合には、AIML-ap の処理系を用いて出力を決定し、その出力を NAOqi フレームワークを介して音声化して会話相手に伝える。あるいは、CACTS を用いた会話タスクのスケジューリングが行われる際には、スケジューリング結果に合わせて会話外行動実行モジュールは発話を行い、AIML-ap の処理系も更新された、あるいはされなかった会話タスクを参照し、次の発話を行う。

5.2 実験 1 : CACTS の客観的な印象の評価

本節では、CACTS を用いたロボットの切り替えの様子の客観的な印象を得るための実験の概要とその結果の議論について述べる。

5.2.1 目的

サービスロボットは周囲に様々な人が存在するような、公共空間での活動が期待される。従って、ロボットのコミュニケーションでは Goffman の聞き手の分類 [29] における傍観者が存在しており、ロボットの挙動は傍観者のロボットへの印象形成にも関わってくるため、客観的な印象についても評価する必要がある。そこで、実験 1 では、ロボットが会話外から割り込みを受けるというシチュエーションを再現した動画を被験者に見せて印象評価を行なうことで CACTS を用いたロボットの切り替えの様子の客観的な印象を調査する。

5.2.2 実施内容

実験には、ロボットが会話タスクスケジューリング手法 CACTS に従って動作する場合と、比較として常に最新の会話タスクを優先するというスケジューリングルールに従う場合の 2 つの動画を用意した。比較に用いたスケジューリングルールは、Schegloff による考察 [37] に基づいている。以下では、比較に用いたスケジューリングルールをベースライン手法と呼称する。被験者にはそれぞれのロボットの印象、及び、どちらのロボットが好ましく感じたかを尋ねた。ロボットの印象は 1 から 7 の数字で評価される。それぞれのスケールには両極な意味の形容詞対が置かれている。印象の項目は礼儀正しさ、自然さ、柔軟さ、賢さ、適切さ、話してみたいと感じたか、の 6 つとした。最初の二項目は社交性、次の二項目は知性、最後の二項目はロボットの親しみやすさを代表するものとして選んだ。実験に用いた 4 つの会話タスクを表 5.2 に示した。会話タスクの優先順位づけに関わる因子と割り込む会話タスク、そして割り込む先の会話タスクとの優先度の差を一覧にした。また、図 5.2 には、それぞれのビデオでの会話タスクのタイムラインを示した。各会話タスクが実行されている間は濃色、中断中は淡色で表現している。予備実験では会話切り替え全体の印象から実装への知見を得るため、「割り込まれた直後」「隣接ペアの拡張形の完成直後」「会話タスクの終了後」の 3 つのパターンが入るように会話タスクを設定した。特に、愚痴 (Emotion が Anger かつ目的なし) の会話タスクは、4 章の実験ではクレームとして見なされてしまい適切な結果が得られなかったため、実験内で切り替えが妥当であったかを検証する。

5.2.3 結果

実験には、18 人の日本の大学生 (女性:2, 男性:16; 平均年齢 : 20.1; SD = 0.62) が参加した。Figure 5.3 はそれぞれのポリシの印象の値を示しており、柔軟さと賢さの項目で有意な差がみられた。これは、我々の提案した手法がベースライン手法より知的に見えたということを示している一方、人間らしさやロボットの親しみやすさには大きな影響を与えなかったことが明らかとなった。さらに、ロボットの好ましさについての質問でも、10 人が提案手法を 6 人がベースライン手法を好ましく感じたと回答し、2 名がわからないを選択しており、提案手法が好ましく感じられた割合 (55.6%) は 50% から大きく離れるものではなかった。これは被験者は当事者ではないために不当に待たされたというような印象を抱きづらく、提案手法とベースライン手法で差

表 5.2: 実験で使用した会話タスクの一覧

| | Topic | 目的の有無 | 会話の長さ | Emotion | 割り込む会話タスク | 人間関係 | 優先度の差 |
|---|-------|-------|-------|---------|-----------|------|-------|
| A | 健康診断 | 目的あり | 長い | Neutral | — | 同格 | — |
| B | 愚痴 | 目的なし | 長い | Anger | (A) | 同格 | -8 |
| C | 予定確認 | 目的あり | 短い | Neutral | (A) | 同格 | +5 |
| D | 天気 | 目的なし | 短い | Neutral | (B) | 下位 | +1 |

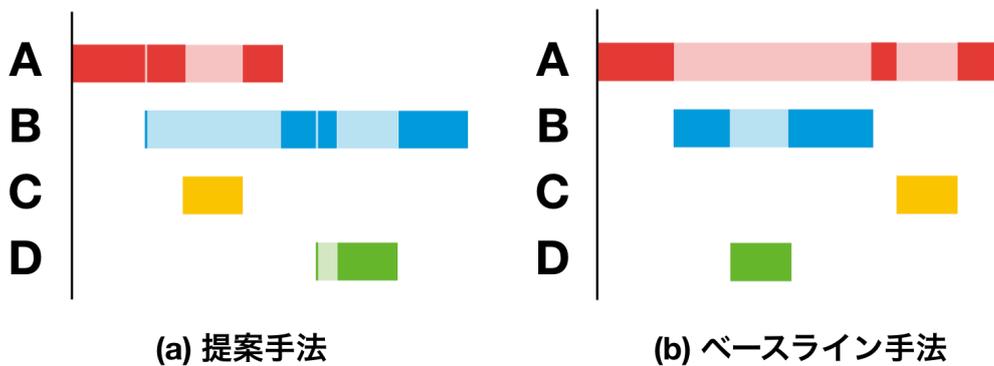


図 5.2: ビデオ内で会話タスクが実行されている時間

を感じづらかったのではないかと考えられる．評価に有意差が現れなかったその他の理由について以下で考察する．

社交性

いくつかのコメントで、「全ての切り替え時に『少々お待ち下さい』...が発生するので、会話の自然さが損なわれる。」などの会話の切り替え時に関する不自然な振る舞いを指摘があった．従って、人間らしさに関して良い結果が得られなかった理由として、社交性はさまざまな具体的な行動に大きく影響されるのではないかと、ということが考えられる．

ロボットへの親密さ

自由記述では2つの理由が言及されていた．一つ目は提案手法が継続的な意図推定をしなかったことである．何人かの被験者はロボットが会話を切り替えすぎるという点について「話し始めたらずっと会話を続けてほしい」、「最小限の会話の切り替えで人を不快にさせないことが大事」と問題視するコメントをしており、会話の切り替え頻度は会話相手の印象に影響を与えていると考えられる．二つ目の理由は、ロボット自身の意図が明確ではなかったということである．ある被験者はベースライン手法の方が望ましいと答えた理由として「会話の切り替え点が明確だったため」と記述していた．複雑な切り替えの決定手法は熟考された行動になる一方、決定がなされた理由は複雑で不明確になってしまう．ロボットの意図が不明瞭だとロボットが予期しな

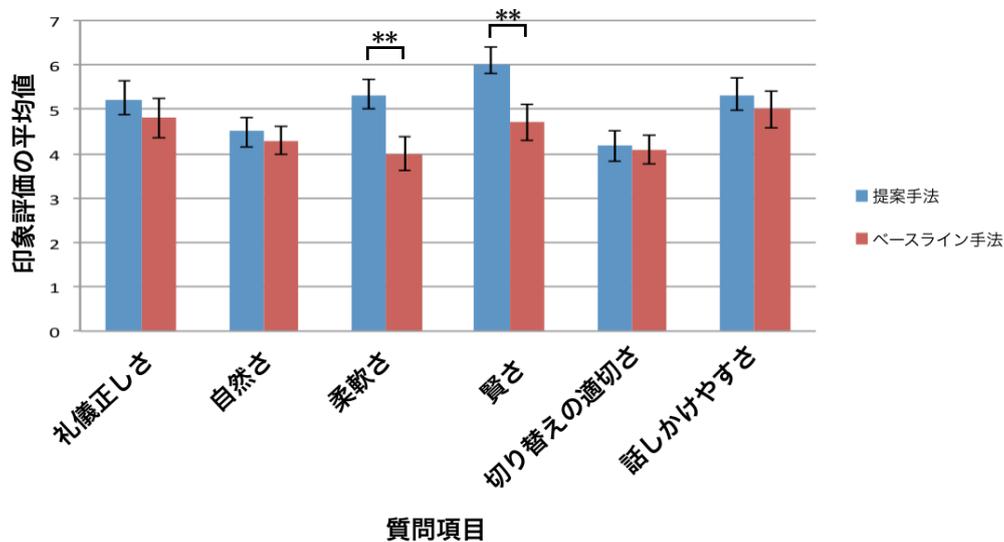


図 5.3: ビデオ視聴後の各手法に対する印象評価の平均値 (**: $p < 0.01$)

い動作をされると考えられてしまうために、そのようなロボットに対しての印象が悪化してしまうと考えられる。そこで、この問題を解決する 2 つの具体例について以下に示す。

因子の削減 単純な手法はロボットの行動の透明性を実現すると考えられる。

理由の説明 会話の切り替えタイミングの決定手法ではないが、ロボットの意図を直接伝えることで不明確さを払拭するできると考えられる。

会話の切り替えタイミングに関する因子は 4 章で検討しており、提案手法において採用した因子については全て影響力が存在することについて議論済みである。因子を減らす選択肢は切り替え自体の適切さを損なう可能性があるため、好ましい手法とは言えない。切り替えの理由の説明を行うと、提案手法の適切さを損なうことなくロボットの意図を明確にすることができる。説明内容をどのようにすべきかは明確ではないが、影響力の強い因子ほど相手に納得されやすいと考えられるので、優先する側に優位な因子のうち最も影響力の高い因子について理由として述べる、といった方法が考えられる。しかし、人間関係や感情について触れることはセンシティブな話題なので、説明の因子として用いることができるのは目的の有無と会話の長さの 2 項目だと考えられる。

5.2.4 まとめ

本実験では、客観的な立場では、提案手法はベースライン手法よりも知的に見えるが、人間らしさやロボットの親しみやすさ、好ましさへの影響はほとんど見られなかった。これは、客観的立場では会話に待たされたというような意識が抱きづらいためだと考えられる。一方、自由記述やアンケートでは切り替えたタイミングそのものよりも、切り替えを行う際のロボットの具体的な行動が意識されていることがわかった。以下に実験 1 で得た知見をまとめる。

- ロボットが同じセリフを繰り返す，などの切り替えに関わる細かい動作がロボットの印象に無視できない影響を与える可能性がある
- 一つの会話タスク中での会話タスクの切り替え頻度が，話し相手の印象に影響する可能性がある
- ロボット自身の意図が明快である方が好まれやすい可能性がある

5.3 実験 2：CACTS の割り込まれる側からの印象の評価

実験 1 では，被験者はロボットが割り込みに対応している様子を収めた動画を視聴したのみで実際に体験はしていなかった．本実験では被験者に割り込まれる側の立場で前節で詳細を述べたロボットとの会話の中で割り込みの対応を体験してもらい，ロボットの個々の割り込みでの対応および会話タスクのスケジューリング手法自体への印象についてアンケートを実施する．アンケート結果とインタラクション中の被験者の振る舞いをもとに，会話タスクのスケジューリング手法 CACTS および会話タスク実行システムについてその有効性を評価する．本実験でも，比較対象として 5 章の予備実験と同様に，常に最新の会話を優先するというスケジューリングルールを採用したシステムを用意し，同様に割り込みの対応を体験してもらった．比較用のスケジューリングルールは以下ではベースライン手法と呼称する

5.3.1 実験概要

被験者はまず実験室で実験の趣旨について説明を受ける．ロボットや情報技術に対する経験に関する事前アンケートを受けたのち，前半と後半で手法を入れ替え 3 つずつ，計 6 つの会話を行う．各会話タスクごとに被験者にはロールが割り当てられる．被験者はそのロールになりきり，ロール内で指定された目的を達成するためにロボットと会話する．ロールは A5 用紙一枚のサイズに印刷されて被験者に渡される．ロールの 1 例について図 5.4 に示す．被験者はロボットとの会話中に同じくロールが割り当てられた実験者によって 1 度割り込みを受ける．ただし，実験者と被験者の間の人間関係のデータは，現実と同じになるようにした．被験者は，被験者および実験者の会話完了したのち実験中のロボットの割り込みの対応についてのアンケートを回答する．さらに，前半・後半が終了するごとに 3 回の割り込まれた体験を通しての感想についてのアンケートを回答する．最後に，前半と後半でどちらの切り替え手法の方が好ましかったについてのアンケートを回答する．

5.3.2 システムの修正

実験 1 で現れた改善点として，会話を切り替える理由を説明する，というものがあつた．そこで実験 2 では，シナリオ外行動実行モジュールに，すぐに会話を切り替える際には「目的の有無」か「会話の長さ」に差があれば，定型文で説明する発話を出力する機能を追加した．

5.3.3 会話シナリオ

実験に当たって，割り込まれる側の会話タスクは「目的有り」で会話の長さが「長い」ものに限定した．理由として，まず，「目的が無い」会話タスクに対して「目的が有る」会話タスクが割り込む場合，「目的の有無」は優先度に対して最も強く影響力を与える因子のため割り込み側が優先される可能性が高く手法全体の検証には適さないこと，そして目的が無い会話に対して目的がない会話が割り込むというケースは，現実には割り込みではなく雑談への合流である可能性が高く現実的な設定とは言えないことから，割り込まれる側は「目的有

- ・ 場所 : 職場
- ・ TODO : コピー機のトラブルシューティング
- ・ 質問ワード : コピー機が壊れた
- ・ 周囲の人 : 同僚(同期)が仕事をしている
- ・ その他の状況 :
 - ・ A4の書類を印刷しようとコピー機を使用している時にエラーに遭遇
 - ・ 用紙がありません[エラーコード : C003] と表示されたので、コピー機のカセットを確認したが、A4の紙は入っていた。
 - ・ 何度か試しても同じエラーがでてしまった。また、何度も試している時に普段とは違うガシャガシャとした音がした。

図 5.4: 実験で使用されたロールの一例

り」に限定した。加えて「短い」会話タスクは実行される時間が短くなるので、その間に割り込みが起こることは稀だと考えられるため、「長い」会話タスクに限定した。

一方、割り込む側の会話タスクは、「目的の有無・会話の長さ・割り込み主の感情」の3つの因子が「目的有り・短い・Neutral」と「目的有り・長い・Anger」と「目的なし・短い・Neutral」の3つに限定した。また、提案手法を用いた場合の会話の切り替えタイミングについて表 5.3 に示した。なお、表 5.3 では実験中に被験者の感情が大きく動く可能性は低いと考えられるため被験者の感情は Neutral であると仮定した結果となっている。実験では、表の項目それぞれに対応する会話シナリオを2つづ用意しており、被験者と実験者の関係と対応する3種類のシナリオをランダムな順番で体験する。会話シナリオの内容については表 5.4 に示した。

また、全ての割り込みシチュエーションにおいて割り込みが起こるタイミングは、隣接ペアの拡張形の中で最初に当たる発話がなされた直後とし、実験者は事前に決定したタイミングで割り込みを行う。これは、隣接ペアの拡張形が完成後に会話を切り替える場合に、割り込み側に切り替わるまでの時間を長くすることで、なるべく割り込まれてすぐに会話を切り替える場合と差別化をするためである。

5.3.4 アンケート

実験中に計3種類のアンケートを行う。被験者がシナリオを1つ終えるごとに行うアンケートでは、そのシナリオで発生した割り込みに対するロボットの対応についての印象について評価するため、ロボットの対応に

表 5.3: 提案手法を用いた場合の各割り込みシチュエーションでの切り替えタイミング

| 被験者から見た 実験者の 上下関係 | 割り込む会話タスクの目的の有無・長さ・感情 高い <—優先順位—> 低い | | |
|-------------------------|---|---------------------|-----------------|
| | 目的有り・短い・Neutral | 目的有り・長い・Anger | 目的なし・短い・Neutral |
| 上位 | 割り込まれた直後 | 割り込まれた直後 | 会話タスクが終了した直後 |
| 同格 | 割り込まれた直後 | 隣接ペアの拡張形が 完成した直後 | 会話タスクが終了した直後 |
| 下位 | 隣接ペアの拡張形が 完成した直後 | 会話タスクが終了した直後 | 会話タスクが終了した直後 |

表 5.4: 提案手法を用いた場合の各割り込みシチュエーションでの切り替えタイミング

| シチュエーション | シナリオ内容 |
|-----------------|--|
| 目的有り・短い・Neutral | オフィスにて、ロボットで午後の商談に使う部屋の確保をしている時に、 ロボットで自分の予定を確認したい人が現れる。 |
| | 従兄弟の家で、料理をしようとロボットとレシピの相談をしていると、 従兄弟が帰ってきてロボットでメールチェックをしたいと言ってくる。 |
| 目的有り・長い・Anger | オフィスにてコピー機が壊れたので、ロボットに相談していたところ、 社内で使用してるメッセージアプリの不具合の報告したい人が現れる。 |
| | 翌日のスキー旅行に備えて準備の確認をロボットに行ってもらっている時、 従兄弟がシミのある服を見つけてロボットにシミ抜きの方法を聞きにくる。 |
| 目的なし・短い・Neutral | オフィスで今日の進捗報告をロボットにしていると、 近くで雑談をしていた人がロボットにテレビドラマの評判をたずねてくる。 |
| | キッチンでの掃除をするため、掃除の仕方をロボットに相談していると、 近くでご飯の話をしていた従兄弟がロボットに昨日の夕飯をたずねてくる。 |

対する印象として「適切さ」「納得感」「丁寧さ」の項目の5段階評価を回答してもらった。また、被験者自身の割り込みシチュエーションに対する理解を確認するために、割り込む側の「会話の必要性」「予想した会話の長さ」「体感としての会話の長さ」「割り込み側の感情」「自分が良いと思う会話の切り替えタイミング」についての質問と、さらに、ロボットの行為が割り込んだ人に対する印象にも影響を及ぼすかを評価するため、割り込んだ人に対する「印象の良さ」「礼儀正しさ」の項目についても5段階で評価を回答してもらった。被験者がシナリオを3つ終えるごとに行うアンケートでは、3回の割り込まれた体験を通しての印象を調査するため、「適切さ」「納得感」「丁寧さ」「自然さ」「柔軟さ」「賢さ」の項目の5段階評価を回答してもらう。全てのシナリオを終えた後に行うアンケートでは、前半と後半のどちらの会話タスクスケジューリング手法を好ましく感じたかと、そしてその理由について回答してもらう。

5.3.5 実験時の問題点

実験中、ロボットの動作についていくつか問題となる点があった。1つは話し相手の顔を見る機能で、割り込まれて以降、ロボットは会話を再開してから元々話していた相手と新しくやってきた人との間をチラチラと交互に見る動作が多発した。実験後のコメントでも多数言及されており、システムに対する悪印象を与える原因となってしまったと考えられる。システムでは、声の向いた方向に向くことでカメラ内に現在話すべき人の顔を認識しその顔を追従するようにしていたが、実際には2人の距離が離れていなかったために元の顔も認識したままになってしまい、このような挙動になってしまったと考えられる。人の位置情報についてより正確な管理が必要だった。また、話し相手の顔をうまく見ることができなかったために、表情認識の機能についても実験者の意図しない表情を認識する場合があった。そこで、実験の初期段階で表情の検出機能をせず1回目の会話タスクと2回目の会話タスクで実験で想定している感情が検出されたとみなして割り込みの処理を行うようにした。感情認識については、多人数会話の場合必ずしも全ての人の表情をカメラできない場合が考えられるので、声の韻律などの情報を用いて、よりロバストな認識を行うことが必要であると考えられる。最後に、ロボットの音声認識の精度の問題があり、何度も聞き手が同じことを言い直さなければならない場面が何度も存在し、自由記述でも認識率の低さについて言及されているものがあった。

5.3.6 結果と考察

実験には、20人の日本の学生(女性:4, 男性:16; 平均年齢: 21.45; SD = 1.54)が参加した。また、実験内で割り込みを行うロールを担当した実験者は日本人の学生(男性; 満21歳)である。実験において、実験者が被験者よりも人間関係が上位となった実験が7件、同格となった実験が6件、下位となった実験が7件であった。図5.5は実験中の様子である。



図 5.5: 3回のインタラクション後の各手法に対する印象評価の平均値 (***: $p < 0.005$)

提案手法の好ましさ

アンケートの結果、提案手法が好ましいと答えた割合が0.75、ベースライン手法が好ましいと答えた割合が0.05、どちらも変わらないと答えた割合は0.20だった。提案手法とベースライン手法で差がないと仮定した

場合とのカイ二乗検定は $p < 0.05$ で有意に差が認められた。また、提案手法を先に実験した被験者は 11 名、ベースライン手法を先に実験した被験者は 9 名だったが、実験順による有意差は見られなかった。次に、手法ごとの印象評価の結果について図 5.6 に示す。提案手法の方が割り込みに対する対応の適切さ、納得感の項目で有意に高い値になっており、提案手法のほうが好ましいというアンケートの結果と一致している。また、実験者と被験者の人間関係ごとのそれぞれの手法の印象評価の結果を図 5.7 に示した。全体での結果に基づき片側検定を行なった結果、(b) の同格の「適切さ」、(c) の下位の「適切さ」「納得感」の項目で有意な差が認められた。また、「適切さ」「納得感」については有意差は出なかったが、平均値の値は提案手法の方が高くなっており、人間関係に関係なく提案手法の方が好ましいと考えられていると言える。

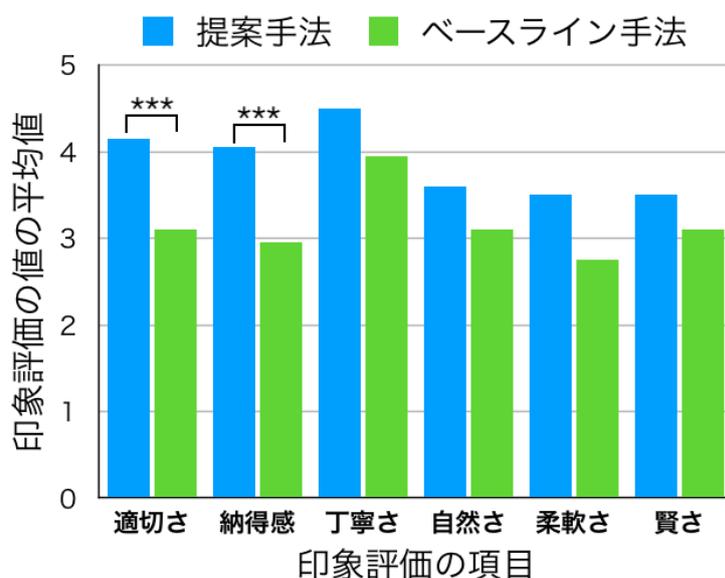
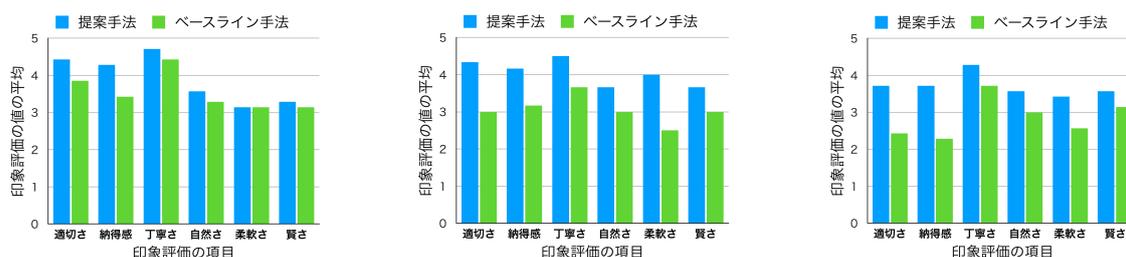


図 5.6: 3 回のインタラクション後の各手法に対する印象評価の平均値 (***: $p < 0.005$)



(a) 実験者が被験者より上位の場合 (b) 実験者と被験者が同格の場合 (c) 実験者が被験者より下位の場合

図 5.7: 実験者と被験者の人間関係別の 3 回のインタラクション後の各手法に対する印象評価の平均値 (\dagger : $p < 0.05$; 片側検定)

個々のシチュエーションの会話の切り替えタイミング

次にそれぞれのシチュエーションでの提案手法の振る舞いの適切さについて議論する。まず、被験者が考えた会話の切り替えタイミングとして適切だと考えられるタイミングについて、実験者が被験者より上位となった実験のアンケート結果を図 5.8、実験者が被験者と同格だった実験のアンケート結果を図 5.9、実験者が被験者より下位となった実験のアンケート結果を図 5.10 に示す。なお、アンケートの「自分が良いと思う会話の切り替えタイミング」の項目の選択肢は「ロボットが切り替えたタイミング」「割り込まれた直後」「会話に一区切りがついた時」「会話が終わった後」「その他(自由記述)」となっており、図 5.8、5.9、5.10 では、ロボットが切り替えたタイミングと答えた者については、それぞれの実験で実際にロボットが会話を切り替えたタイミングに置き換えている。また、それぞれのシチュエーションについて、実験者が想定した会話タスクの優先順位づけに関わる因子の水準を用いてシステムが決定した会話の切り替えタイミングと、被験者が判断した会話タスクの優先順位づけに関わる因子の水準で修正してシステムが決定した会話の切り替えタイミングが、アンケートの「自分が良いと思う会話の切り替えタイミング」とどれだけ一致したかについて、実験者が被験者より上位だった場合を図 5.11、実験者が被験者と同格だった場合を図 5.12、実験者が被験者より下位だった場合を図 5.13 に示す。この際、被験者が選択した「会話に一区切りがついた時」とシステムが切り替えた「隣接ペアの拡張系が完成した後」は、一致しているか不明なため、図の中では緑で分けて示した。

図 5.8、5.9、5.10 のように、被験者にとって好ましいと考える会話の切り替えタイミングはばらつきがあることがわかる。一方で、被験者による各シチュエーションの会話タスクの優先順位づけの因子の判断を用いて提案手法による会話の切り替えタイミングの判断を行なった場合、図 5.11、5.12、5.13 のように、実験者の想定した各シチュエーションの会話タスクの優先順位づけの因子を用いた場合より一致率が高くなることはなかった。ただし、有意差は見られない。従って、会話タスクの切り替えタイミングはシチュエーションのみによって定まるものではなく、割り込まれた側の個性に依る部分がある可能性が考えられる。

個人によってシチュエーションごとの割り込みに対する対応の判断は提案手法の判断した会話の切り替えタイミングは、「割り込まれた直後」「会話に一区切りがついた時」「会話が終わった後」という単位では、図 5.9(b)を除き、被験者にとって好ましいと考える会話の切り替えタイミングのうち最も割合の高いタイミングとなっており、提案手法の会話タスクの切り替えタイミングの判断は一定の妥当性があると考えられる。一方、図 5.9(b)で「会話に一区切りがついた時」が好ましいと思われたタイミングが低かった原因として、コピー機が壊れたシチュエーションでは「プリンターの故障は比較的緊急性が高いのでこちらのほうが優先されるべきであると思った」や、旅行の準備をしているシチュエーションでは「しみ抜きは緊急性が高いので、直ぐに割り込み会話に入るのが適切であると思った。」というコメントがあり、「緊急性が高い」話題について先に優先すべきではないかという判断がなされた可能性が考えられる。「緊急度」は 5 章の実験の結果では切り替えタイミングに与える影響力が小さい因子だったため、除外したが、考慮が必要だった可能性が考えられる。また、このような結果になった理由として、長期的な人間関係の視点が存在していた可能性が考えられる。被験者の考える会話の切り替えタイミングを「会話が終わった後」だとした理由として「人間関係にその後問題が出る可能性はあると思ったが、自分の方を優先してくれたほうがよいと感じた」というコメントがあった。感情をあらわにしている人と対応する場合には、長期的な関係性など、その場の会話だけにとどまらない要因が絡んでくることが示唆されている。

もう一点、図 5.10(a)では「ロボットが切り替えたタイミング」を選んだ被験者はおらず、「会話に一区切りがついた時」を選択していることから、ここで被験者が想定している会話の切り替えタイミングは「隣接ペアの拡張系が完成した後」ではない別のタイミングではないかと考えられる。その他の項目では、商談の場所を

確保するシナリオにおいて「俺が話してるからまってくれ」、料理の相談をするシナリオにおいては「そこまで話したら自分の会話の後でよくない？ってタイミングで切り替えられたのでちょっと違和感を感じた」と、会話の切り替えタイミングがキリが悪く感じられているような意見がみられた。切り替えタイミングとして好ましくないと感じた理由として、待っている間の暇さがあると考えられる。参考になる意見として、キッチンの掃除について相談するシナリオでは、掃除のために「用意するものを教えてもらった後に、用意している間に割り込み対応すればいいのでは」ないか、コピー機が壊れたシナリオでは「「あなたは」荷物の確認に一度行って下さい」など、指示を出してくれると自然に感じそう」というように、待ってもらう間に作業を行わせると良いのではないかという意見があった。隣接ペアの概念とは全く異なるが、途中で会話を切り替える場合にはロボットが待ってもらう間にやってもらう作業が見つかるように会話を誘導することで、キリの悪さを解消できる可能性がある。

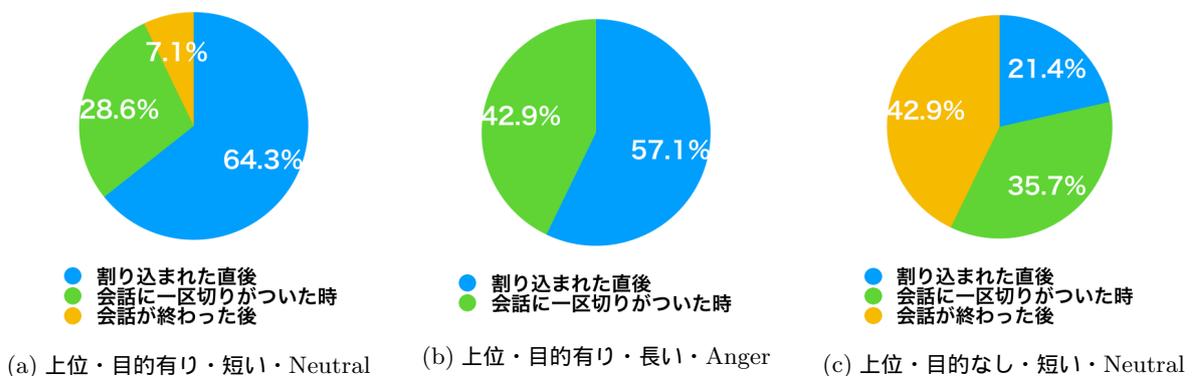


図 5.8: 実験者が人間関係で上位の時の各シチュエーションで好ましいと考える会話の切り替えタイミング

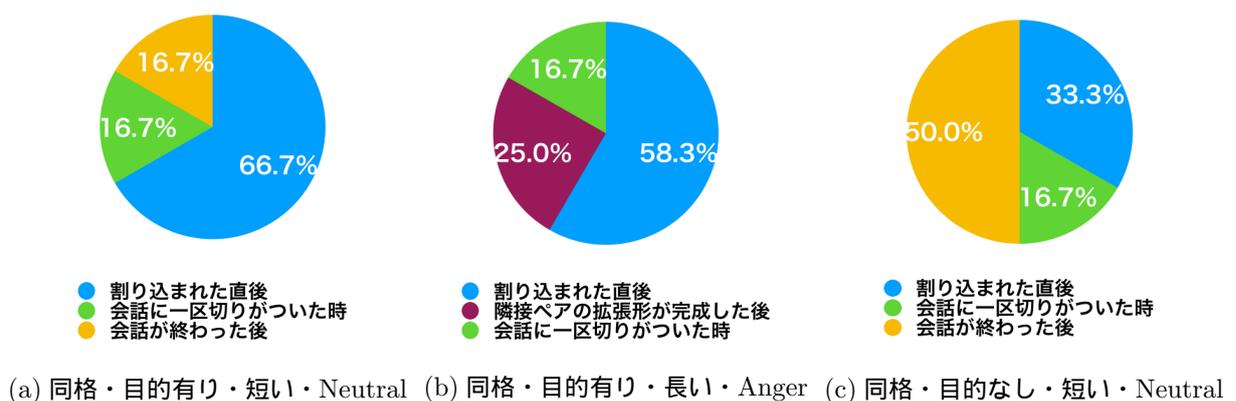
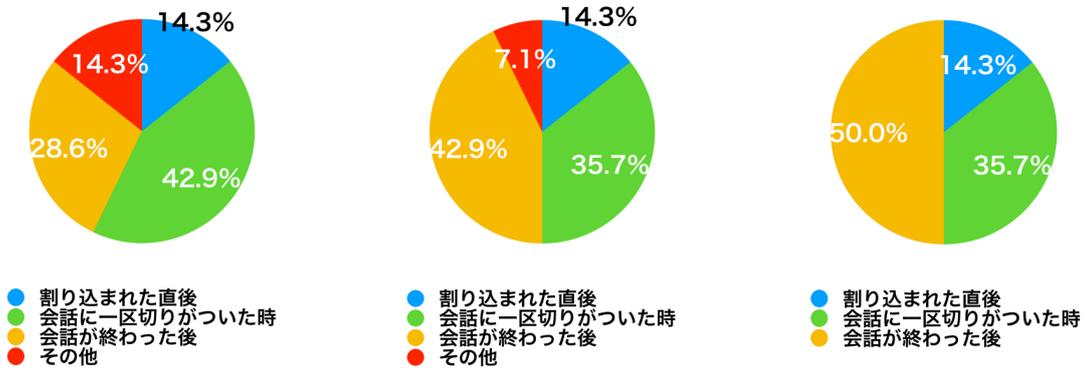


図 5.9: 実験者が人間関係で同格の時の各シチュエーションで好ましいと考える会話の切り替えタイミング

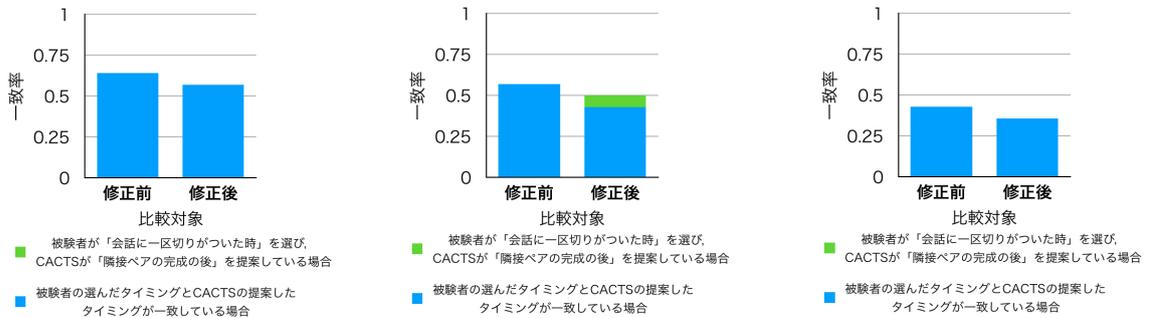
個々のシチュエーションでのロボットへの印象

各シチュエーションでのロボットへの印象について、実験者が被験者より上位となった実験のアンケート結果を図 5.14、実験者が被験者と同格だった実験のアンケート結果を図 5.15、実験者が被験者より下位となった実験のアンケート結果を図 5.16 に示した。まず、実験者が被験者より上位となった実験では「上位・目的



(a) 下位・目的有り・短い・Neutral (b) 下位・目的有り・長い・Anger (c) 下位・目的なし・短い・Neutral

図 5.10: 実験者が人間関係で下位の時の各シチュエーションで好ましいと考える会話の切り替えタイミング

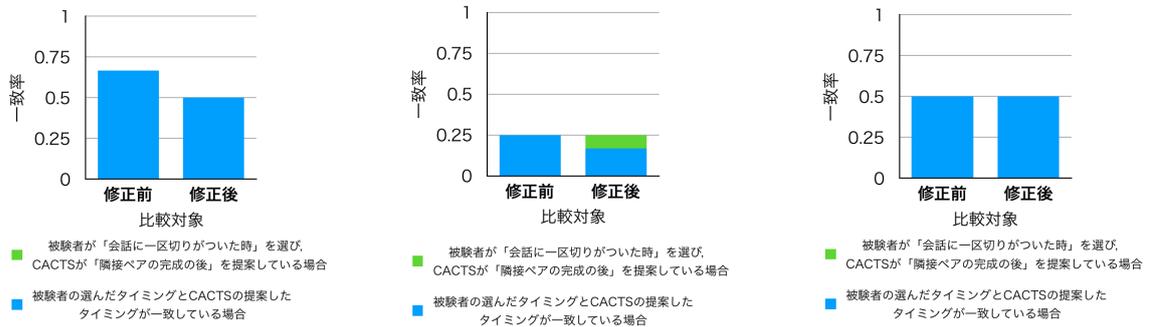


(a) 上位・目的有り・短い・Neutral (b) 上位・目的有り・長い・Anger (c) 上位・目的なし・短い・Neutral

図 5.11: 実験者が人間関係で上位の時の各シチュエーションで好ましいと考える会話の切り替えタイミングと提案手法の一致率

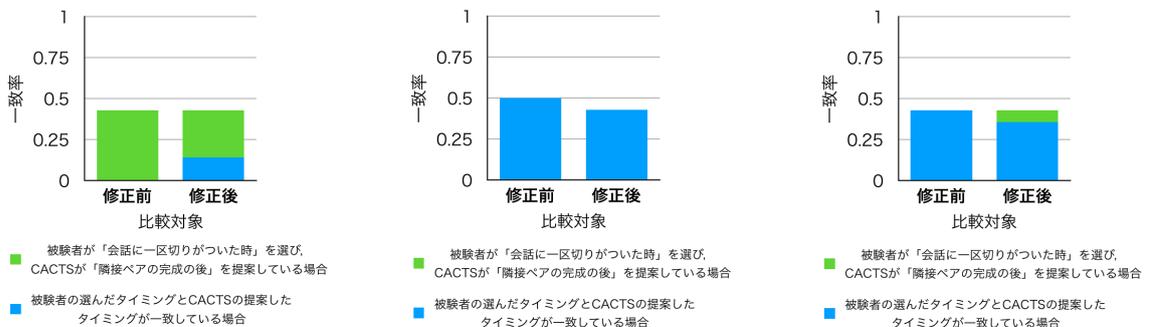
有り・短い・Neutral」と「上位・目的有り・長い・Anger」の切り替えるパターンは提案手法とベースライン手法で同じなので、手法間での有意な差などは得られなかった。一方、「上位・目的なし・短い・Neutral」のケースでは適切さの項目で提案手法の方が好ましいとする有意差が見られた。また、被験者が考える会話の切り替えタイミングも、提案手法と同じタイミングが最も高い割合となっている。

次に、実験者が被験者より上位となった実験では「上位・目的有り・短い・Neutral」のケースで納得感と自然さの項目で有意な差が見られた。このケースでは切り替えるパターンは提案手法とベースライン手法で同じである。提案手法とベースライン手法の相違点としては、提案手法の場合には切り替え時の理由として「”すぐ終わるので”少々お待ちください」という言葉を付け足して会話の切り替えを行なっている。ただし、この対応は実験者が被験者より上位となった実験の「上位・目的有り・短い・Neutral」でも行われているが、こちらの場合ではほとんど差が見られない。このようになった理由として、「上位・目的有り・短い・Neutral」となったケースでは、同格であった場合と比べて人間関係の差の分だけ優先度が高くなっているため、わざわざ理由を聞くまでもなく納得できたが、「同格・目的有り・短い・Neutral」の場合では、優先度の差がそれほど大きくないため、理由の説明がない場合には割り込んだ側が優先されたことに対する納得感が得られなかったのではないかと考えられる。一方、「上位・目的有り・長い・Anger」と「上位・目的なし・短



(a) 同格・目的有り・短い・Neutral (b) 同格・目的有り・長い・Anger (c) 同格・目的なし・短い・Neutral

図 5.12: 実験者が人間関係で同格の時の各シチュエーションで好ましいと考える会話の切り替えタイミングと提案手法の一致率

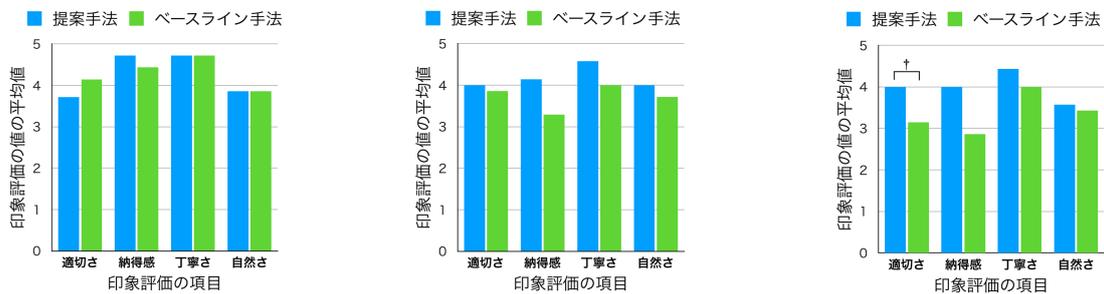


(a) 下位・目的有り・短い・Neutral (b) 下位・目的有り・長い・Anger (c) 下位・目的なし・短い・Neutral

図 5.13: 実験者が人間関係で下位の時の各シチュエーションで好ましいと考える会話の切り替えタイミングと提案手法の一致率

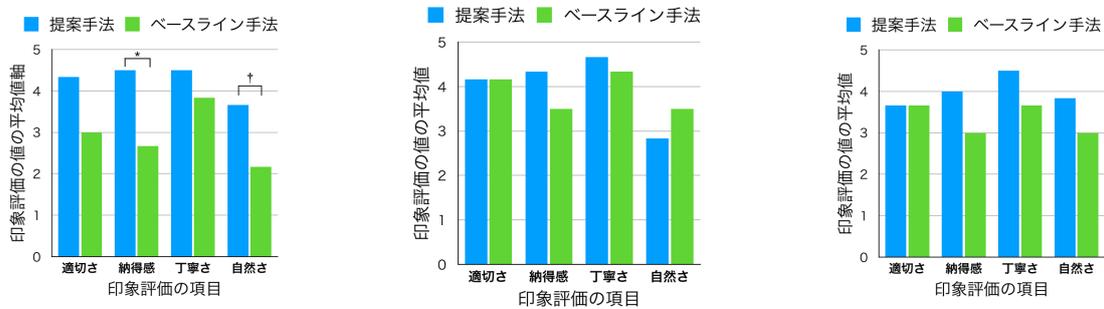
「Neutral」のケースでは、ベースライン手法と提案手法で切り替えたタイミングが異なるにも関わらず、印象評価の項目について有意な差は見られなかった。全体としては提案手法の方が良いという結果を得たので、提案手法が平均値で優っている項目については母数が小さいために有意差が出なかっただけでも考えられるが、「同格・目的有り・長い・Anger」のケースについては自然さの項目ではベースライン手法の方が優っており、明らかに他とは異なる傾向を示している。これは前述したように、緊急度や長期的な人間関係という視点での判断が影響しているためと考えられる。

最後に、実験者が被験者より下位となった実験では、「上位・目的有り・短い・Neutral」のシチュエーションでベースライン手法の方が適切さ、納得感の項目の平均値が高くなっている。これは前述したように、提案手法が会話の切り替えタイミングとして判断した「隣接ペアの拡張形が完成した後」のタイミングに違和感を覚えていたためだと考えられる。「上位・目的有り・長い・Anger」のケースでは全ての項目で有意な差があり、「上位・目的なし・短い・Neutral」のケースでも自然さの項目で有意な差があったので、提案手法の切り替えタイミングの方が好ましく感じられていたと考えられる。



(a) 上位・目的有り・短い・Neutral (b) 上位・目的有り・長い・Anger (c) 上位・目的なし・短い・Neutral

図 5.14: 実験者が人間関係で上位だった場合の印象評価の結果 (†: $p < 0.05$; 片側検定)

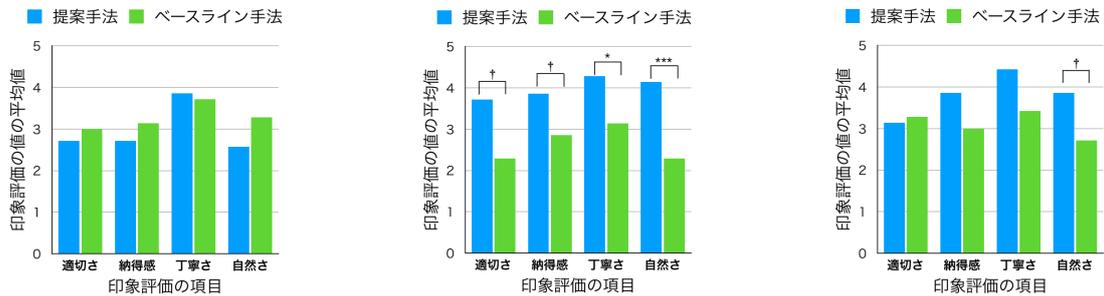


(a) 同格・目的有り・短い・Neutral (b) 同格・目的有り・長い・Anger (c) 同格・目的なし・短い・Neutral

図 5.15: 実験者が人間関係で同格だった場合の印象評価の結果 (*: $p < 0.05$, †: $p < 0.05$; 片側検定)

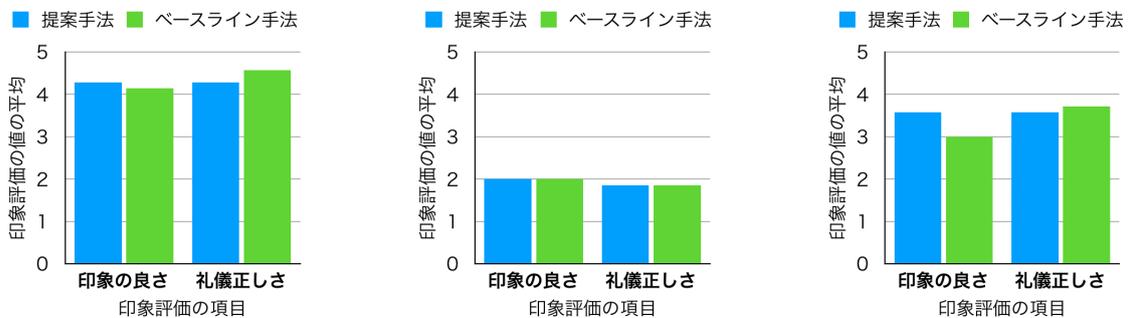
割り込んだ人に対する印象

割り込んできた相手に対する印象を、実験者と被験者の関係ごとに図 5.17, 5.18, 5.19 に示した。多くのシチュエーションで提案手法の方が割り込んだ人への印象が良くなりやすいという傾向が見られたが、有意な差は認められなかった。傾向として現れていることから、今後母数を増やして検証すべき点だと考えられる。



(a) 下位・目的有り・短い・Neutral (b) 下位・目的有り・長い・Anger (c) 下位・目的なし・短い・Neutral

図 5.16: 実験者が人間関係で下位だった場合の印象評価の結果 (*: $p < 0.05$, ***: $p < 0.005$, †: $p < 0.05$; 片側検定)



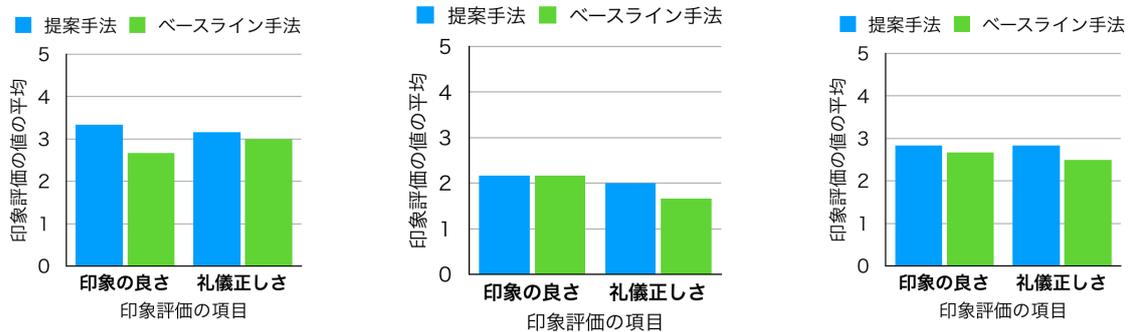
(a) 上位・目的有り・短い・Neutral (b) 上位・目的有り・長い・Anger (c) 上位・目的なし・短い・Neutral

図 5.17: 割り込んだ側が人間関係で上位だった場合の割り込んだ人への印象評価の結果

自由記述の分析

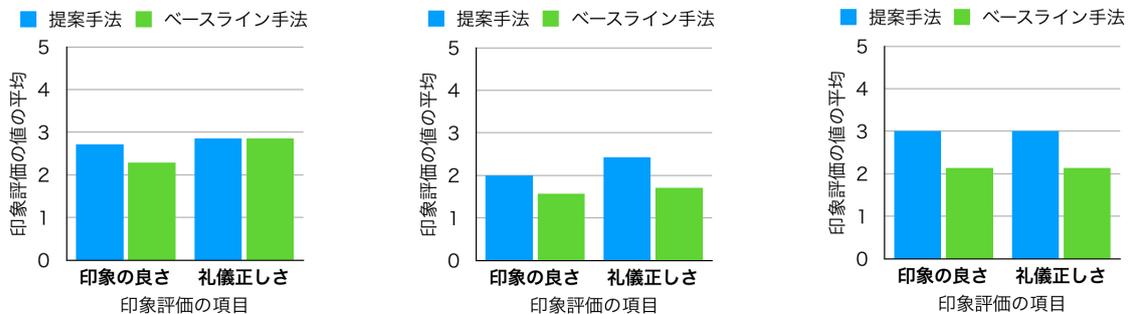
アンケートの自由記述でより好ましい手法を判断した時の理由について尋ねたところ、提案手法を選んだケースでは、優先順位づけに関するものとして「会話の重要性とか考えている気がした。」「優先度が低い要件をわりこませなかった」などが挙げられており、そのほかには、「会話を切り替えるときの対応が丁寧に感じたため。」「すぐ終わるので」という一言がよかったから。」など、細かい動作に関して好ましいと感じたもの、「後半の方がよりスムーズに会話が成立しているように感じた。」「しっかりと確認作業をして、スムーズに進んでいたから」というシステムのユーザビリティに関するものがあった。また、どちらの手法も変わらないとしたものでは、「特にこれといって具体的な違いが分からなかった。」「全会話の違いがよくわからん」といった、シチュエーションの違いが感じられなかったというものと、「タイミングが悪いと感じたときには、「会話のコンテンツ価値判断/人間関係への配慮」どちらかが欠けていたように感じた」、「どちらとも、適切だと思う場合とそうでない場合があった」というように、それぞれに問題があったとするものがあった。ベースライン手法の方を選択したケースでは、理由として「割り込みに対する返答の順序」の正しさ、があげられていた。

また、各手法ごとのロボットについての感想として、会話の切り替えを現在の会話者に確認していない、という点をあげたものが複数見られた。会話の優先度が明らかな場合に尋ねることは会話のスムーズさを損ねることになるためロボットの印象を下げることに繋がるが、差が小さいと考えられる場合には割り込まれた人の



(a) 同格・目的有り・短い・Neutral (b) 同格・目的有り・長い・Anger (c) 同格・目的なし・短い・Neutral

図 5.18: 割り込んだ側が人間関係で同格だった場合の割り込んだ人への印象評価の結果



(a) 下位・目的有り・短い・Neutral (b) 下位・目的有り・長い・Anger (c) 下位・目的なし・短い・Neutral

図 5.19: 割り込んだ側が人間関係で下位だった場合の割り込んだ人への印象評価の結果

意思を確認することは有効だと考えられる。また、家庭でのシチュエーションの例ではロールでははっきりしなかったロボットの所有者についての言及があった。CACTS では、人間同士の関係性を因子として用いたが、シチュエーションによってはロボットと人間との関係性についても考慮する必要性が示唆されていた。

5.4 まとめ

本実験の結果として、傍観者などの第三者的な立場から見た場合、会話の切り替えタイミングそのものよりも会話の切り替える時の行動の方が印象に影響を与えていることがわかった。また、割り込まれた側にとっては、常に最新の会話タスクを優先するベースラインシステムよりも、会話の構造や会話の優先度を考慮したCACTSのほうが好ましいことがわかった。一方で、次にあげるCACTSが対象としなかった因子が会話の切り替えタイミングに影響を与えている可能性が示された。

- 割り込まれた回数・頻度
- タスクの緊急度
- ロボットと人間との関係性
- 長期的な人間関係の維持

また、CACTS は割り込まれた側の考える会話タスクの切り替えタイミングは、会話タスクの目的の有無や長さ、人間関係、割り込み主の感情といった、外部の要因によって定まるという前提に立っていたが、実際には割り込まれた側がどのようなタイミングで切り替えれば良いと思うかには個人差があり、必ずしもシチュエーションによって一つの答えになるとは定まらなかった。そこで、実験中の意見で示された方針としては、

- 割り込まれた側と会話を切り替えていいかについて相談する
- 会話を切り替えるタイミングとして適切となるように会話を誘導する

というものがあつた。人間同士の場合に、CACTS のような割り込まれてすぐに会話を切り替える手法、割り込まれた側と相談する手法、途中で中断できるように会話を誘導する手法はそれぞれが実際に使われているのか？また、使われているとしたらどのように使い分けられているのかを今後明らかにする必要があると考えられる。

第 6 章

結論

本章では、これまでの議論を総括し、今後の課題について述べる。

6.1 総括

本研究では、将来予想されている労働者不足問題に際して活躍の期待されているロボットが直面する課題として、それまで交わしてきた会話とは全く異なる内容について話しかけられる可能性を挙げ、この会話外からの割り込みに対応する手法について提案を行った。提案手法 CACTS (Context Aware Conversation Task Scheduling) は、会話のみで完結するタスクを対象に、割り込みによって生じた会話タスクに切り替えるタイミングをスケジューリングする。会話タスクの切り替え可能なタイミングの候補は、会話内の最小のインタラクション構造である隣接ペアとその拡張形に注目し、隣接ペアの拡張形が完成する前に会話タスクの切り替えが起こることは望ましくないことを実験で示し、割り込まれた直後か、隣接ペアの拡張形の完成後、あるいは会話の終了後のいずれかに定めた。また、会話の切り替えを行うかの判断の基準としては、「達成すべき目的の有無」「会話の長さ」「話者同士の人間関係」「話者の感情」の 4 つが影響力を持つことを実験より明らかにした。CACTS はこの 4 つの因子を元に会話の切り替えタイミングを決定する。

1 つ目の実験では、動画の視聴によって傍観者視点でのロボットの切り替えの様子の印象を評価し、傍観者の視点では具体的な会話の切り替えタイミングよりも具体的な行動の方が印象に影響を与えていることがわかった。また、被験者に CACTS を実装したロボットと実際にインタラクションをしてもらった実験では、ロールプレイで割り込みのシチュエーションを再現した際のロボットに対して受ける印象について評価し、常に最新の会話タスクに切り替える場合に比べ、好ましい印象を与えることを示した。従って、提案手法 CACTS は会話外からの割り込みに対する会話タスクのスケジューリング手法として一定の有効性があると結論づける。一方で、割り込まれた側が会話タスクの切り替えタイミングとして好ましいと考えるタイミングには個人差があり、CACTS のような割り込まれたシチュエーションによって会話タスクの切り替えタイミングを一意に定める手法には限界があることも示された。

6.2 今後の課題と展望

本実験では割り込む側の視点での評価が行えなかった。今後現実的环境下で実験を行い、割り込む側の印象についても評価していきたいと考えられる。

本研究において会話タスクは、会話タスクを記述するためのスクリプトによって静的に記述されていた。会

話を切り替えるタイミングを決定するための指標として用いた隣接ペアは、コーパスのアノテーション情報として付与されていたり [40]、秋葉らの研究 [38] のように手書きのルールでの検出が行われており、自然言語処理の手法での検出を行うための環境が整いつつあり、CACTS は隣接ペアの自動判定によって、動的な発話生成の手法への応用が可能となると考えられる。CACTS で用いた因子は 2 値で表現できるもの (会話の目的の有無・会話の長さ・怒っているか) や、3 値に表現できるもの (人間関係) であるので、応用が容易であると考えられる。

また、割り込みのシチュエーションに対して一意に会話タスクの切り替えタイミングを切り替えることの限界も実験内で示された。CACTS に限らず、割り込まれた側と相談する手法、途中で中断できるように会話を誘導する手法を組み合わせることを考えた場合、実際の人同士での割り込みの様子について分析し、割り込みに対する対応方法がどのように使い分けられているのかを知る必要がある。人同士のコミュニケーションの観察に基づいたロボット実装としては、社会学のエスノメソドロジーの手法を用いた秋谷らの介護施設での会話開始部分に注目した研究 [27] や、Ghosh らの美術館でガイドを行うロボット [48] などがあり、人同士の行動を分析することで、より好ましい行動をロボットが獲得できるようになると考えられる。

さらに、本研究では会話タスクのスケジューリングをメインで扱ったが、実際には割り込み自体の検出や、割り込みを円滑に処理するための行動など、様々な関連研究が存在し、それらはまだ未開拓な部分も多い。割り込みの検出技術としては、周辺で行われる多数の発話に対してロボットの応答義務を推定する杉山らの研究 [49] がある。また、割り込みを円滑に処理するための行動方針としては、一方の会話に対応する間にもう一方を待たせるという状況は、ポライトネスの概念における相手に悪印象を与えるフェイスの侵害行為に当たると考えられ、それらの緩和手法について扱うポライトネス研究 [50] などが参考になると考えられる。

謝辞

本研究を進めるにあたって、ご指導頂きました慶應義塾大学環境情報学部准教授高汐一紀博士に深く感謝致します。また、慶應義塾大学環境情報学部教授徳田英幸博士には、本論文の執筆に当たって御助言を賜りました事を深く感謝致します。慶應義塾大学高汐研究室の同期、後輩方には折りに振れ貴重なご助言を頂きました。長時間にわたる実験にも快く協力くださった学友に深く感謝します。実験にここに深く感謝の意を表します。最後に、大学3年間に渡る生活を支えてくれた家族に感謝致します。

23 January 2018

Takumi Horie

参考文献

- [1] Dan Bohus and Eric Horvitz. Models for multiparty engagement in open-world dialog. In *Proceedings of the SIGDIAL 2009 Conference: The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pp. 225–234. Association for Computational Linguistics, 2009.
- [2] Ronald P. A. Petrick and Mary Ellen Foster. Planning for social interaction in a robot bartender domain. In *In Proceedings of ICAPS 2013, Special Track on Novel Applications*, pp. 389–397, 2013.
- [3] Emanuel A Schegloff and Harvey Sacks. Opening up closings. *Semiotica*, Vol. 8, No. 4, pp. 289–327, 1973.
- [4] Emanuel A Schegloff. *Sequence organization in interaction: Volume 1: A primer in conversation analysis*, Vol. 1. Cambridge University Press, 2007.
- [5] 内閣府. 平成 29 年版高齢社会白書. 行政, 2017.
- [6] 総務省. 平成 29 年版情報通信白書. <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h21/index.html>, 2009.
- [7] Iolanda Leite, Carlos Martinho, and Ana Paiva. Social robots for long-term interaction: a survey. *International Journal of Social Robotics*, Vol. 5, No. 2, pp. 291–308, 2013.
- [8] Joseph Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Commun. ACM*, Vol. 9, No. 1, pp. 36–45, January 1966.
- [9] Terry Winograd. Procedures as a representation for data in a computer program for understanding natural language. Technical report, MASSACHUSETTS INST OF TECH CAMBRIDGE PROJECT MAC, 1971.
- [10] 河原達也ほか. 音声対話システムの進化と淘汰: 歴史と最近の技術動向 (特集: 音声対話システムの実用化に向けて). *人工知能学会誌*, Vol. 28, No. 1, pp. 45–51, 2013.
- [11] Victor Zue, Stephanie Seneff, and James Glass. Speech database development at mit: Timit and beyond. *Speech Communication*, Vol. 9, No. 4, pp. 351–356, 1990.
- [12] 竹林洋一. 音声自由対話システム tosburg; cd02d36. gif; ユーザ中心のマルチモーダルインタフェースの実現に向けて. *電子情報通信学会論文誌 D*, Vol. 77, No. 8, pp. 1417–1428, 1994.
- [13] 神田直之, 武田龍, 大淵康成ほか. Deep neural network に基づく日本語音声認識の基礎評価. 研究報告 音声言語情報処理 (SLP), Vol. 2013, No. 8, pp. 1–6, 2013.
- [14] 山本幹雄, 伊藤敏彦, 肥田野勝, 中川聖一ほか. 人間の理解手法を用いたロバストな音声対話システム. *情報処理学会論文誌*, Vol. 37, No. 4, pp. 471–482, 1996.
- [15] 駒谷和範, 河原達也, 清田陽司, 黒橋禎夫ほか. 柔軟な言語モデルとマッチングを用いた音声によるレストラン検索システム. 情報処理学会研究報告 音声言語情報処理 (SLP), Vol. 2001, No. 123 (2001-SLP-039),

- pp. 177–182, 2001.
- [16] 吉村貴克, 徳田恵一, 益子貴史, 小林隆夫, 北村正. Hmm に基づく音声合成におけるスペクトル・ピッチ・継続長の同時モデル化. 電子情報通信学会論文誌 D, Vol. 83, No. 11, pp. 2099–2107, 2000.
 - [17] キャンベルニック, ブラックアラン. Chatr: 自然音声波形接続型任意音声合成システム. 電子情報通信学会技術研究報告. SP, 音声, Vol. 96, No. 39, pp. 45–52, 1996.
 - [18]
 - [19] 大須賀智子, 堀内靖雄, 西田昌史, 市川薫. 音声対話での話者交替/継続の予測における韻律情報の有効性. 人工知能学会論文誌, Vol. 21, No. 1, pp. 1–8, 2006.
 - [20] 北村裕貴, 湯浅将英, 武川直樹. 「人の視線」と「ロボットの視線」が伝える意図の比較分析: ロボットの「コチラガホシイ」の視線動作デザイン (一般, 「コミュニケーション研究の今, その先」及び一般). 電子情報通信学会技術研究報告. HCS, ヒューマンコミュニケーション基礎, Vol. 109, No. 457, pp. 55–60, 2010.
 - [21] 橋本卓弥, 平松幸男, 辻俊明, 小林宏. ロボット受付嬢 saya を用いたリアルなうなづきに関する研究. 日本機械学会論文集 C 編, Vol. 73, No. 735, pp. 3046–3054, 2007.
 - [22] 小林一樹, 山田誠二. 擬人化したモーションによるロボットのマインド表出. 人工知能学会論文誌, Vol. 21, No. 4, pp. 380–387, 2006.
 - [23] 神田崇行, 石黒浩, 小野哲雄, 今井倫太, 前田武志, 中津良平. 研究用プラットフォームとしての日常活動型ロボット. 電子情報通信学会論文誌 D, Vol. 85, No. 4, pp. 380–389, 2002.
 - [24] 光永法明, 宮下善太, 宮下敬宏, 石黒浩, 萩田紀博. コミュニケーションロボット robovie-iv の開発とオフィス環境での日常対話. 日本ロボット学会誌, Vol. 25, No. 6, pp. 822–833, 2007.
 - [25] 神田崇行, 今井倫太, 小野哲雄, 石黒浩ほか. 人-ロボット相互作用における身体動作の数値解析. 情報処理学会論文誌, Vol. 44, No. 11, pp. 2699–2709, 2003.
 - [26] 松坂要佐, 東條剛史, 小林哲則. グループ会話に参与する対話ロボットの構築. 電子情報通信学会論文誌 D, Vol. 84, No. 6, pp. 898–908, 2001.
 - [27] 秋谷直矩, 丹羽仁史, 岡田真依, 山崎敬一, 小林貴訓, 久野義徳, 山崎晶子ほか. 高齢者介護施設におけるコミュニケーションチャンネル確立過程の分析と支援システムの提案. 情報処理学会論文誌, Vol. 50, No. 1, pp. 302–313, 2009.
 - [28] 高梨克也ほか. 会話構造理解のための分析単位: 参与構造 (i 連載チュートリアル i , 多人数インタラクションの分析手法 [第 6 回]). 人工知能学会誌, Vol. 23, No. 4, pp. 538–544, 2008.
 - [29] Erving Goffman. *Forms of talk*. University of Pennsylvania Press, 1981.
 - [30] Harvey Sacks, Emanuel A Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction*, pp. 7–55. Elsevier, 1978.
 - [31] 桂子藤井. 発話の重なりについて: 分類の試み. 言語文化と日本語教育, Vol. 10, pp. 13–23, dec 1995.
 - [32] 長谷川紀子. 日本語学習者の割り込み発話. 千葉大学日本文化論叢, No. 6, pp. 105–90, 2005.
 - [33] 堂坂浩二, 島津明. 協調的な話し言葉生成. 電子情報通信学会技術研究報告. NLC, 言語理解とコミュニケーション, Vol. 96, No. 294, pp. 9–16, 1996.
 - [34] 高梨克也. 複数の焦点のある相互行為場面における活動の割り込みの分析 (i 特集 i , 相互作用のマルチモーダル分析). 社会言語科学, Vol. 14, No. 1, pp. 48–60, 2011.
 - [35] 高橋裕己, C Kennington, 船越孝太郎, 中野幹生, 菅野重樹. 確率的ルールを用いた多人数ロボット対話.

- 人工知能学会第 28 回全国大会論文集 2M5-OS-20b-4, 2014.
- [36] 石川真也, 船越孝太郎, 篠田浩一, 中野幹生. 多人数対話ロボットの実現にむけたマルチモーダル対話データの収集と分析. 人工知能学会第 27 回全国大会論文集 1K3-OS-17a-5, 2013.
- [37] Emmanuel A Schegloff. Opening sequencing. *Perpetual contact: Mobile communication, private talk, public performance*, pp. 326–385, 2002.
- [38] 秋葉巖, 松山洋一, 小林哲則ほか. 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. 研究報告音声言語情報処理 (SLP), Vol. 2013, No. 10, pp. 1–8, 2013.
- [39] Alexander Shyrovkov, Andrew Kun, and Peter Heeman. Experimental modeling of human-human multi-threaded dialogues in the presence of a manual-visual task. *Proceedings of 8th SIGdial*, pp. 190–193, 2007.
- [40] 吉田奈央, 高梨克也, 伝康晴. 対話におけるあいづち表現の認定とその問題点について. 言語処理学会第 15 回年次大会発表論文集, pp. 430–433, 2009.
- [41] 伝康晴ほか. 会話・対話・談話研究のための分析単位: 隣接ペア (i 連載チュートリアルj 多人数インタラクションの分析手法 [第 4 回]). 人工知能学会誌, Vol. 23, No. 2, pp. 271–276, 2008.
- [42] Simon Keizer, Milica Gasic, François Mairesse, Blaise Thomson, Kai Yu, and Steve Young. Modelling user behaviour in the his-pomdp dialogue manager. In *Spoken Language Technology Workshop, 2008. SLT 2008. IEEE*, pp. 121–124. IEEE, 2008.
- [43] 中西亮輔, 井上昂治, 中村静, 高梨克也, 河原達也ほか. 円滑な発話権制御のための談話行為の連鎖に基づくフィラーの生起と形態の予測. *SIG-SLUD*, Vol. 5, No. 03, pp. 18–24, 2017.
- [44] Richard Wallace. The elements of aiml style. *Alice AI Foundation*, 2003.
- [45] André Neves and Flávia Barros. Xbotml: a markup language for human computer interaction via chatterbots. *Web Engineering*, pp. 329–346, 2003.
- [46] Clement Creusot. Actbot: Sharing high-level robot ai scripts. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*, pp. 172–178. IEEE, 2016.
- [47] ST Hawk, J Van der Schalk, and AH Fischer. Moving faces, looking places: The amsterdam dynamic facial expressions set (adfes). In *12th European Conference on Facial Expressions, Geneva, Switzerland, 2008*.
- [48] Madhumita Ghosh and Hideaki Kuzuoka. An ethnomethodological study of a museum guide robot 's attempt at engagement and disengagement. *Journal of Robotics*, Vol. 2014, , 2014.
- [49] 杉山貴昭, 船越孝太郎, 中野幹生, 駒谷和範. 多人数対話におけるロボットの応答義務の推定. 人工知能学会第 29 回全国大会, 2015.
- [50] 宇佐美まゆみ. ポライトネス理論研究のフロンティア: ポライトネス理論研究の課題とディスコース・ポライトネス理論 (i 特集j 敬語研究のフロンティア). 社会言語科学, Vol. 11, No. 1, pp. 4–22, 2008.