

会話参加状態を考慮した振る舞いをするロボットのシステムアーキテクチャ

A System Architecture for Conversation Robot Considering Participant Structure

菅原一真*1 浅野秀平*1 赤川優斗*1 藤江真也*2 小川哲司*1 小林哲則*1
Kazuma Sugawara Shuhei Asano Yuto Akagawa Shinya Fujie Tetsuji Ogawa Tetsunori Kobayashi

*1早稲田大学 Waseda University *2千葉工業大学 Chiba Institute of Technology

This paper proposes participants management rules for conversation robots. These rules define the methods of how to participate into and leave from conversation. These methods manage the participation structure of conversation. This paper also proposes a conversation robot system architecture that enables various kind of conversation robots to behave the same function according to the proposed rules. In this architecture, behavior command generating module generates abstract action command according to the proposed rules. By converting this abstract action command to hardware-dependent action command, robots which have different physical characteristics can behave to perform the same function of the rules. To clarify the effectiveness of the proposed rule and architecture, in addition, we implement two conversation robot which include 'SCHEMA' and 'Sota'.

1. はじめに

多人数会話への参加, 離脱を中心に, 参与構造形成のために必要な会話プロトコル (= 円滑な会話の成立のために守るべき手続きの約束ごと) の精緻化について検討する. また, そこで会話ロボットがとるべき行動の記述を, 行動が担うプロトコル上の機能レベルでの表現を介したものにすることで, 身体性の異なる複数のロボットへの適用を容易にすることを試みる.

多人数会話においては, それぞれの人は会話に加わったり抜けたりすることで, 会話参加者は動的に変わる. また, 会話に形式的に参加しながらも, 中心となる話の展開には集中せず, 他の人と話を始めることも多い. 円滑に会話を進めるためには, 会話参加者は, こうした複雑な状態を把握するとともに, それぞれの状態に応じた振る舞いの規則を適切に用意しておくことが重要である.

これまでに, 多人数会話における動的な会話参加者の入れ替わりを扱ういくつかの研究がなされている. 坊農らは, 学会でのポスター発表における会話を対象として, 会話参加者の視線や身体位置等による身体動作レベルの会話参加手続きについて分析的な検討を行っている [坊農 2004]. Bohus らは, 人の主観的および客観的参加状態と次にとるべき行動で構成される会話参与状態モデルを提案し, 会話参加者が動的に変化する一対多の会話システムを実現している [Bohus 2009a], [Bohus 2009b]. 山際らは, 会話参加者の参加態度を推定しながら身体位置を調整する機能を会話ロボットに実装している [山際 2013]. 筆者らも, 会話ロボットが従うべきプロトコルを総合的に整理する中で, 会話ロボットが自らの参与構造の理解状態を表出することの必要について指摘し, 会話参加者が動的に変化する状況において, 身体の向きを用いてそれを表現する機能を実装してきた [Kobayashi 2010], [Kobayashi 2013].

これらの研究に共通する点は, 会話参加者の変化を扱うために必要な行為が, 身体動作の現象レベルで記述されていることである. このとき, ある状況でとるべきそのロボットなりの振る舞いは, 対象のハードウェアに依存した形で記述されることになる. このため, ある会話ロボットに対する設計された振

る舞いを, 身体性の異なるロボットに利用することは難しい. Pepper や Sota といった身体性の異なる様々な会話ロボットが, 日常生活に入り込んで活躍の場を得つつあるが, 会話ロボットとしてどのような振る舞いをさせるべきかについては, 他のロボットでの知見を利用することなくゼロから設計しているのが現状である.

本研究では, 人同士が円滑な会話を行うために必然性を持って行っている振る舞いを, 単なる現象レベルでなく, 機能・役割レベルで再考する. 機能レベルに相当する上位の行動記述と, 具体的な身体動作に相当する下位の行動記述を分離することで, 身体性が異なる場合も上位の部分は共通化し, ハードウェアに依存する部分を下位に隠蔽することができる. このことによって, 設計者は下位の行動のみを設計すれば, 他ロボットで検討された知見を再利用して, 重要な点を外さずに会話ロボットを実現できることが期待できる.

本稿では, 人同士が会話参加者の入れ替わりを扱うために従っている約束事を, 既に会話参加している人とこれから参加したい人とが行う, 抽象的な要求の発信とその応答によって定義するとともに, それらの扱いに適した会話ロボットのアーキテクチャを提案する.

2. 参与役割の拡張

2.1 従来の参与役割の考え方

Goffman や Clark が提唱する会話参与構造では, 会話に対して承認された参与者 (ratified participant) とそれ以外の立ち聞き者 (overhearer) に分け, 参与者については話し手 (speaker), 受け手 (addressee), 傍参与者 (side-participant), 立ち聞き者については傍観者 (bystander) をなどがある. これらは主に, 話し手の発話調整のために分類されたものである. 本節では, 会話参加, 離脱の現象をパターンとして挙げ, そこで必要となる参与役割の検討を行う. 以降, 承認された参与者を指して参加者と呼ぶ.

2.2 会話への参加・離脱のパターン

2.2.1 傍観者からアプローチして参加するパターン

傍観者が会話にアプローチして, 参加者に受諾, または拒否されるパターンを図 1 に示す.

B : (近寄ってくる)
P1 : 「最近何か映画見た？」
P2 : 「いや見てないー」
B : 「テレビでアナと雪の女王やってたけどそれは見た？」
(傍観者からのアプローチ)

P3 : 「ああ、見たよー」(受諾)
or
P3 : 「今違う話してるから待ってて」(拒否)

図 1: 傍観者からのアプローチする例。P1~P3 は会話参加者, B は傍観者を表す。

P1 : 「最近何か映画見た？」
P2 : 「いや見てないー」
P3 : 「テレビでアナと雪の女王やってたけどそれは見た？」
B : (近寄って P3 の話に頷く)
P3 : 「B 君は見た？」(参加者からのアプローチ)
P3 : 「ああ、見たよー」(受諾)
or
P3 : 「ごめん、聞いているだけだよ...」(拒否)

図 2: 参加者からアプローチする例。P1~P3, B は図 1 と同様。

傍観者は、発話できるタイミングで一時的に発話権を取り発話する。これに対して参加者が応答すれば、傍観者は参加者として認められたことになる。一方、参加を制止するような応答があれば、参加者にはなれない。

2.2.2 参加者からアプローチして参加するパターン

参加者から傍観者を参加させるパターンを図 2 に示す。

参加者から傍観者に対して話しかけ、それに傍観者が応答する。傍観者が応答すれば参加者となり、そうでなければ傍観者のままだい。参加者からのアプローチは発話を伴うため、これを実行できるのは発話権を持つ話し手に限られる。

2.2.3 話し手以外の参加者から参加を促すパターン

発話を伴わなくとも、傍観者に顔を向ける、位置をずらして参加する場所を空けるなど、参加者が傍観者の参加を促すことがある。この行動自体は傍観者の役割を変えるものではないが、二つの効果がある。

一つ目は、話し手へのアプローチの促進である。参加を希望する傍観者の存在や、参加者自身がその傍観者を参加させたいことを、話し手に知らせることで、2.2.2 のパターンになることを期待するものである。二つ目は、傍観者の参加意欲の確認である。特に、参加者の行動に対して傍観者が拒否をした場合、会話を聞いていただけということが明確になる。

2.2.4 話し手が参加者に集中を促すパターン

参加者が会話とは別の事象に注意を向けることがある。例えば、スマートフォンの操作、別の会話グループへの参加などが挙げられる。このような行動を取る人に対し、話し手が会話へ集中を戻すことを促す会話パターンを図 3 に示す。

話し手で発話することで集中を戻すためには、呼びかけ発話が必ず必要となる。一方、集中を失っていても参加者である限り発話をする権利はあるので、自ら集中を戻して会話に戻る場合は文脈に沿った内容を発話してよい(特に断りは必要ない)。

2.3 役割の拡張

これまで述べた各パターンを踏まえて、参与役割に以下を加えて再整理する。

P1 : 「最近何か映画見た？」
P2 : 「いや見てないー」
I : (スマートフォンを操作して話を聞いていない)
P1 : 「ねえねえ I 君、何か見た？」(参加者からのアプローチ)
I : 「ああ、アナと雪の女王を見たよー」(受諾)
or
I : 「ごめん、ちょっと待って...」(拒否)

図 3: 参加者からアプローチする例。P1, P2 は会話参加者, I は会話へ集中していない人を表す。

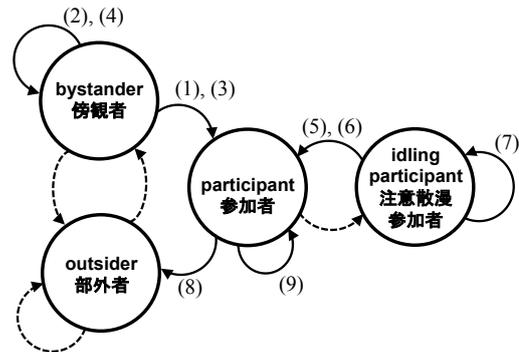


図 4: 会話参加状態遷移モデル

idling participant (注意散漫参加者) 従来の傍観者のうち、参加している会話とは別の事象に注意を向けている者。会話参加はしているが、会話内容に注意を向けていないため、他の参加者が受け手にしたい場合、注意を戻す行動をとらなければならない。

outsider (部外者) 従来の参加者でも立ち聞き者でもない者。会話参加意欲もなく、参加者認証もされていないため、従来の参与役割では対象外とされていた。

3. 会話参与者管理規則

3.1 会話参加状態遷移モデル

前節で再整理した参与役割を状態とした状態遷移モデルを定義し、本人と会話の参加者間の手続きによって状態が遷移する。この手続きを参加状態管理規則と呼ぶ。規則は、要求と応答の対で構成される。応答は直前の要求を受け取った相手だが、その要求を受諾するか拒否するかのいずれかである。受諾されれば別の状態に遷移し、拒否されればその状態に留まる。

以降、各規則がどのような手続きで構成されるかを、図 4 に示したモデルにしたがって説明する。図中の実線は規則による状態遷移である。点線は、特に規則によらない状態遷移である。例えば、部外者が傍観者になるためには、会話を聞くために場に近づけばよく、参加者から受諾の応答を得る必要はないと考える。また、図では参加者 (speaker, addressee, side-participant) を participant としてひとまとめにしている。これは、参加状態に関しては参加者間の違いを考慮する必要がないためである。

3.2 傍観者が会話に参加するための規則

3.2.1 傍観者からアプローチする規則

図中 (1), (2) の遷移に対応する。いずれも傍観者が確立要求を発信するが、参加者が受諾した場合は (1) 参加者になり、

参加者が拒否した場合は (2) 傍観者に留まる。

確立要求は、傍観者による発話を伴う行動を表す。この発話には、2.2.1 で挙げたような文脈に沿った発話や、「ねえねえ」のように直接参加を要求するようなものが含まれる。

受諾は、確立要求の発話に対する返答や、頷きなど、参加者と認めた行動である。拒否は、「ちょっと待って」などのような明確に参加を拒否する返答を指す。

3.2.2 参加者からアプローチする規則

図中 (3), (4) の遷移に対応する。いずれも参加者が確立要求を発信するが、傍観者が受諾した場合は (3) 参加者になり、傍観者が拒否した場合は (4) 傍観者に留まる。

確立要求は、参加者による傍観者に向けた発話を伴う行動を表す。発話内容は 2.2.2 に挙げたような文脈に沿った発話や、「ちょっといい？」などの直接的な勧誘も含む。なお、発話を伴う行動であるため、参加者の中でも話し手 (speaker) のみが発信できる。

受諾や拒否に関しては、2.2.2 に挙げたように、確立要求の発話に対する返答をすれば受諾であり、「聞いているだけです」などと発話すれば拒否となる。

3.3 注意散漫参加者が注意を戻す規則

3.3.1 注意散漫参加者が自ら戻す規則

図中 (5) の遷移に対応する。注意散漫参加者は、会話の内容を集中して聞いてはいないが、自らが発話をする権利 (話し手になる権利) をもっている。そのため、その権利を行使して発話することが注意散漫参加者の復帰要求となる。

この要求は、現象としては傍参加者が発話権をとつたに過ぎず、他参加者がそれを拒否して注意散漫参加者に留まらせることはできない。したがって、この要求は原則受諾となる。

3.3.2 話し手からアプローチする規則

図中 (6), (7) の遷移に対応する。いずれも、話し手から集中要求を発信するが、注意散漫参加者が受諾すれば (6) 参加者に復帰し、注意散漫参加者が拒否すれば (7) 注意散漫参加者のままとする。

集中要求は、2.2.4 で挙げたような「ねえねえ」などのような、注意を会話に向けるような表現を伴う発話を表す。

受諾は、話し手発話に対して返答するなど受け手としての行動となる。拒否は、単に無視する (集中要求に気づかない場合など) ことや、明示的に「ちょっと待って」など拒否をする発話となる。

3.4 参加者が会話から離脱する規則

図中 (8), (9) の遷移に対応する。会話を離脱しようとする参加者が離脱要求を発信し、他の参加者が受諾すれば (8) 当該参加者は部外者となり、他の参加者が拒否すれば (9) 参加者に留まる。

離脱要求は、明確に「ちょっと抜けます」という発話や、黙ってその場を立ち去るという行動である。受諾は、「分かりました」などの発話や、そのまま会話を続けるといったものである。拒否は、「ちょっと待って」「ひとつだけ確認させて」など、離脱を留まらせる発話をするのである。

4. 参加意欲度

4.1 参加意欲度の必要性

これまで見てきた会話参加状態遷移モデルにおいて傍観者が参加者になるところに注目すると、いずれも発話を伴う要求と応答によって規則が構成されるため、特に参加者側は話し手のみが関わるものとなっている。一方、2.2.3 で見たよう

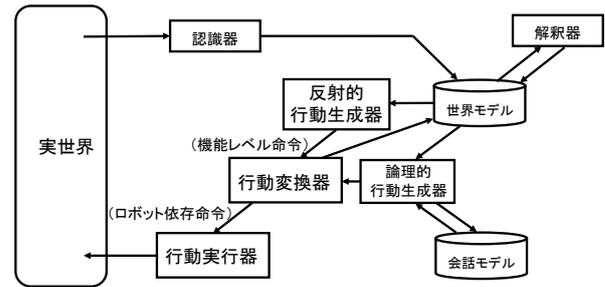


図 5: アーキテクチャ

に、発話を伴わなくとも、他の参加者や傍観者自身の行動によって、他の参加者がある傍観者を参加させたいことをアピールしたり、傍観者自身が参加を希望することを確認することができる。

これらの現象は当該傍観者の役割を変化させるようなものではないものの、話し手が傍観者の存在に気づくことが確立要求を発信するきっかけになるような場合がある。このように、各参加者や傍観者が持つ、当該傍観者を参加させようとする強さを参加意欲度と定義する。以降では、どのような現象がそれらに影響するかについて説明する。

4.2 傍観者の参加意欲度

傍観者が会話に参加したいと思う強さである。傍観者であることは、参加者がその存在を認識している、話を聞いていることも分かっている。これから、ある程度参加意欲が高いことがうかがえるが、本心では聞いていただいだけかもしれない。

傍観者の参加意欲が高いことが確認できる現象としては、3.2.1 で示した確立要求の発信が挙げられる。また、2.2.3 で見たような話し手以外の参加の促しを受け入れるということもあげられる。逆に、話し手以外の参加の促しを拒否した場合は参加意欲が低いことが確認できる。

4.3 参加者の参加意欲度

参加者が、傍観者を参加させたいと思う強さである。傍観者 1 人に対して、全参加者がそれぞれ別の強さをもつ。

全ての参加者が傍観者に気づいているとは限らない。傍観者に気づいていない参加者についてはこの値をそもそも定義することができない。ある参加者が 2.2.3 で見たような参加の促しをすることは、その参加者が傍観者を会話に参加させたいということの現れである。したがって、このような行動をとった場合には、参加者の (その傍観者に対する) 参加意欲度は高いということが確認できる。

5. ロボットシステムアーキテクチャ

5.1 アーキテクチャの概要

提案するロボットシステムアーキテクチャは、会話コンテキスト依存の行動生成器を含む会話タスク依存部分、解釈器や機能レベルでの抽象的な反射行動生成器を含む会話プロトコルに従う抽象的な命令を生成する部分、環境認識器や行動変換器、行動実行器のような実世界部分に大別できる。先述の会話参加者管理規則を定義し、提案アーキテクチャを構成することで会話タスクレベル、会話参加者管理規則レベル、実際の振る舞い生成レベルで実装を分離させることができる。

アーキテクチャの構成を図 5 に示す。アーキテクチャによってロボットが動作するまでの流れは以下の通りである。

- 実会話プロトコルレベルの解釈器が、実世界レベルの認識器による情報を用いて会話参加状態などを解釈する。
- 会話プロトコルレベルの反射行動生成器が、解釈器による情報を元に抽象的なプロトコルレベルの行動命令を生成する。
- 実世界レベルの行動変換器が、プロトコルレベルの反射行動生成器による抽象的な行動命令をロボット毎の具体的な身体動作レベルの行動命令に変換し、行動実行器が逐次実行する。

会話プロトコルレベルでは、会話参与者管理規則における各種要求履歴や各人の参加状態のような概念的な情報を管理する。従って、人の位置情報や顔向きなど具体的な情報を管理する必要がなくなる。実世界レベルでは、会話プロトコルレベルで決定したロボットが発信する抽象的な要求を、行動変換器によって各ロボットの身体性を考慮した実際の要求発信方法を決定する。また、認識器で参与者が発信する要求を認識する。この実装の分離によって、異なる身体性を持つロボットで同じ会話プロトコルを考慮した振る舞いを実現することができる。

5.2 振る舞いを生成する例

異なる身体性を持つロボットが同じ機能の振る舞いを生成する例として、脚部にモータを保持して位置移動できる会話ロボット SCHEMA と、モータを持たず位置移動ができない会話ロボット Sota の2種類のロボットが、傍観者を会話参加するよう誘導する振る舞いをする事例を考える。

位置移動できる SCHEMA は、位置をずれて傍観者が会話参加できる空間を空けるといった行動によって参加を促すことができる。対する Sota は位置移動ができないため、SCHEMA と同じ行動を取ることができない。よって Sota は、腕や首といった脚以外の身体部位を用いて促さなければならない。

この事例で提案アーキテクチャを用いた場合、以下の流れで目標である会話への参加を促すことが達成される。

- プロトコルレベルの反射的行動生成器が、抽象的な機能レベルの行動命令である「会話参加の促し」を生成する。
- 実世界レベルの行動変換器が、ロボット毎に具体的な行動命令に変換する。SCHEMA へは位置移動命令、Sota へは腕、首を動かす命令に変換する。
- 各ロボット毎の行動実行器がそれぞれ変換された具体的な行動命令を実行する。

以上の過程により、会話プロトコルレベルの実装部分は共通のモジュール群を使用して身体性の異なるロボットで同じ機能を表す命令を生成し、実行することができる。

また、会話タスクレベルでは、発話内容を決定する。会話参加者が動的に変化するので会話文脈の管理が必要だが、参加者変化時の行動を考慮する必要がなくなる。よって、会話コンテンツ開発者が振る舞いを考慮しないコンテンツ開発が可能となる。その結果、ロボット開発コストの削減が期待できる。

6. 実装

本研究では、提案アーキテクチャに従い、会話参与者管理規則に従う振る舞いをする会話ロボットの実装を行う。実装対象は会話ロボット SCHEMA と会話ロボット Sota である。

環境認識器として、マイクロソフト社が開発したモーションキャプチャ機器である Kinect を用いる。2台の Kinect で検

知できる範囲内にいる人の顔向き、位置座標を取得し、公開する役割を持つ。

解釈器として会話参加状態検知モジュールを実装した。会話参加状態検知モジュールは、kinect を用いて取得した人の位置、顔向きの情報から、会話参加状態を判定して情報を公開する役割を持つ。

会話参加状態検知モジュールが公開する参加状態を基に、反射的行動生成器が会話参与者管理規則に従う抽象的な命令を生成し、行動変換器が SCHEMA と Sota の身体性の違いを考慮した命令に変換し、各行動実行器が逐次実行する。

7. まとめ

本研究では、会話参加者の入れ替わりを中心とした、参与の状態の動的変化を扱うために必要な約束事を、既に会話参加している人とこれから参加したい人とが行う、抽象的な要求の発信とその応答によって定義するとともに、それらの扱いに適した会話ロボットのアーキテクチャを提案した。また、身体性の異なる2つのロボットに実装し、移植の容易性を検証した。

ここで扱った参加者の変化を扱う問題は、通信におけるプロトコルに例えれば、通信路の確立にあたる。通信においては、データ再送やフロー制御のような信頼性のあるデータ送信、セッション管理や送信するデータの選択など、ここで扱った通信路の確立以外にも様々なレイヤのプロトコルがあるが、それらはそれぞれに会話において同様の機能を見出すことができる。今後こうした問題についても、ここで提案したアプローチを適用し、会話ロボットの従うべき振る舞いの一般化モデルの構築を進めていく予定である。

参考文献

- [山際 2013] 山際康貴, 上原孝紀, 蔵田洋平, 大島直樹, P. Ravindra De Silva, 岡田美智男: 多人数会話に基づくソーシャルインタフェースにおける身体配置の調整について, Human-Agent Interaction シンポジウム 2013 論文集, 70-77(2013).
- [Bohus 2009a] D. Bohus, E. Horvitz: Models for Multiparty Engagement in Open-World Dialog, Conference on The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 225-234(2009).
- [Bohus 2009a] D. Bohus, E. Horvitz: Models for Multiparty Engagement in Open-World Dialog, Conference on The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 225-234(2009).
- [Bohus 2009b] D. Bohus, E. Horvitz: Learning to predict engagement with a spoken dialog system in open-world settings, Conference on The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 225-234(2009).
- [Kobayashi 2010] T. Kobayashi: Conversation robot recognizing and expressing paralinguistic information, invited talk at Workshop on Predictive Models of Human Communication Dynamics, Aug. (2010).
- [Kobayashi 2013] T. Kobayashi, S. Fujie: Conversational robots: An approach to conversation protocol issues that utilizes the paralinguistic information available in a robot-human setting, Acoustical Science and Technology, 34, 64-72(2013).
- [坊農 2004] 坊農真弓, 鈴木紀子, 片桐恭弘: 多人数会話における参与構造分析 インタクション行動から興味対象を抽出する, 認知科学, 11(3), 214-227(2004).
- [Goffman 1981] E. Goffman: Forms of talk, Philadelphia, University of Pennsylvania Press.