

多人数会話の場に基づく ソーシャルインタフェースの提案とその応用

Creature-based Multiple-Party Conversation for Social Interface and Its Application

吉池 佑太^{1*} デシルバ ラビンドラ¹ 岡田 美智男¹
Yuta Yoshiike¹ P. Ravindra De Silva¹ Michio Okada¹

¹ 豊橋技術科学大学 情報・知能工学系

¹ Department of Computer Science and Engineering, Toyohashi University of Technology

Abstract: In this study, we propose a creature-based social interface as an interactive social medium based on the multi-party conversation to broadcast the news and information to users. The interface of MAWARI is consisting of three sociable creatures which are designed with minimal designing mechanism. It is capable of providing the information through the multi-party conversation and it's accommodating for user to get the information with less attempt to the conversation. The idea behind the design is that a user does not need several communication channels to establish a conversation. This helps to reduce the observing workload of the user, because during the conversation the creature just uses vocal interactions mixed with attractive body gestures. One motivation of the current study is to be able to evaluate the workload of the user when interface as an interactive social medium.

1 はじめに

私たちの日常的なコミュニケーションには大きく2つの側面がある。1つは情報を伝達するという側面であり、もう1つは、社会的なつながりを志向するという側面である。例えば、上司から部下への指示などは、言葉の意味を正確に伝えることや正確に解釈することなどの側面が強い。一方で、日常における雑談などは、会話の内容や情報伝達の正確性などよりも、会話それ自体を楽しむといった社会的なつながりを志向する側面が強い。

前者のような側面はインタフェースの枠組みとして適用しやすい。例えば、音声対話インタフェースの研究などでは、人-システム間での発話の意味理解や適切な応答生成などに重きが置かれてきた。しかし、人とロボット・エージェントとの社会的な関わり扱う HRI (Human-Robot Interaction) や HAI (Human-Agent Interaction) といった研究分野が近年広がりを見せている。人と親和性のあるインタフェースを実現するうえでも、人とシステムとの社会的関係についての議論は重要である。



図 1: テーブル上の3つの Mawari とユーザ

そこで、本研究では人とロボットとの社会的な関わり方の1つとして、「3つのロボットたちが会話する場をつくり、その会話に人が参与する」というインタラクションの枠組みを新たに提案する。多人数会話の場に基づくことで、雑談のような場を媒介して情報を提供するシステムが実現出来ると考えられる。このような枠組みを利用したインタフェースを、ここでは「多人数会話の場に基づくソーシャルインタフェース」と

*連絡先: 豊橋技術科学大学大学院 電子・情報工学専攻
〒441-8580 愛知県豊橋市天伯町字雲雀ヶ丘 1-1
F 研究棟 (F-413)
E-mail: yoshiike@icd.cs.tut.ac.jp

呼び、このプラットフォームとして Mawari を構築した (図 1).

本稿では、構築した Mawari のコンセプトとシステム構成について述べ、そのインタフェースとしての応用可能性について整理し、評価実験の結果について報告する。

2 研究背景

2.1 関連研究

私たちが普段家庭や職場などで行っているコミュニケーションは、多人数によるものが多い。そのため、多人数会話に参加するロボット [2][3] やエージェント [4] に関する研究がある。これらの研究では、1つのロボット・エージェントが複数人と関わることに焦点が当てられている。一方で、複数ロボット・エージェントに関する研究は、マルチエージェントシステムと呼ばれる、複数ロボット間の振る舞いやスケジューリングの調整などに焦点が置かれたものが多く、ロボット同士がインタラクションする場に人が参与するという形態の研究はほとんどない。

TalkingEye という社会的なエージェントのシステムでは、マルチエージェントシステムに利用されるメカニズムを社会的な調整行為に応用しており、人と社会的な関係を築くことに有効だと報告されている [6]。また、情報をブロードキャストする情報発信メディアとして、2台のロボットを利用することは有効であるという報告や [5]、複数エージェントを利用してユーザの意思決定を支援することは、ユーザを退屈させないことに有効であるという報告もある [7]。

このように、ロボット・エージェント同士による多人数会話の場を利用することの効果について報告されているものの、その数は少なく、未だ初期段階である。そのため、ロボットのコミュニティに人が参加するという枠組みをどのように作るのか、またどのような効果があるのかについて整理することは、HRI/HAI 研究のみならず、次世代のインタフェースの枠組みとしても新たな視点を提供することにつながるだろう。

2.2 ソーシャルインタフェース

一般にソーシャルインタフェースとは、HCI(Human-Computer Interaction) などの分野において、例えば、マイクロソフトエージェントのような、人間的な特徴を利用したシステムのインタフェースのことを指す場合が多い¹。また、ソーシャルインタフェース (Social Interface) は、人と同様な機能をロボットが利用したシ

ステムであり、遠隔存在感ロボットや博物館ガイドロボットなどであるという解釈もある。

一方で、竹内らは、人とシステムの関係性に着目し、「使う、使われる」ものからお互いに「関わりあう」ものへと変化していることを指摘し、社会的な関係を顕在化させるインタフェースのことをソーシャルインタフェースと呼んでいる [1]。本研究におけるソーシャルインタフェースとはこの考えを基にしたものである。

従来の人とシステムの関係での会話は、お互いに「聞き手」と「話し手」になることによって成り立つ。そのため、どちらかが返答をやめてしまうと会話が成り立たなくなり、この関係性は崩れてしまう。しかし、ロボット同士が「聞き手」「話し手」となっていれば会話の秩序は保たれる。ユーザには会話に参加する自由度が存在し、もし不参加の場合でも会話は継続され、参加を強要されることがない。そのため、ユーザは自由意思に応じて会話に耳を傾けながら、自由に割り込むことや、もしくは別のことに集中することが可能である。

多人数会話の場に基づくソーシャルインタフェースは、このように、会話の場がある程度自律的に維持され、会話参加への自由度を生み出す。このような自由度を「緩やかな共同性」と呼んでいる [8]。

3 Mawari

3.1 コンセプト

私たちの日常では、周囲の人の雑談を耳にすることで、計らずも情報を得ることがある。例えば、電車の中で「今日は雨降るらしいから傘必要だよ」「え、そうなの?今は晴れてるのに」という会話を耳にすれば、今日の天気予報についての情報を得たことになる。もし、その会話をしている人が友人たちであれば、会話に参加することで、より具体的な情報を得ることができるだろう。

Mawari はこのような雑談 (多人数会話) の場をつくり、その場を介してユーザに情報提供を行うソーシャルインタフェースである。会話への参加自由度が存在するため、一对一の音声インタフェースに比べて、ユーザの心的負荷が軽減されると考えられる。

3.2 システム構成

ここでは、Web 上のデータを RSS により取得し、その情報を基にユーザに提供するというシステムとして構築した。

システムは図 1 に示したように、PC と Mawari から構成される。PC にはマイクとスピーカが接続され、音声認識、音声合成、Web 上の情報の取得、および会

¹http://en.wikipedia.org/wiki/Social_interface

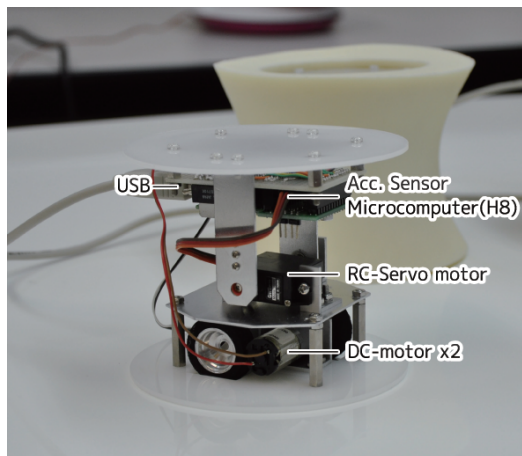


図 2: Mawari のハードウェア概要

話内容を生成する処理を行う。Mawari は USB を介して PC と接続され、制御情報に基づき振る舞いの動作を行う。この構成に基づき、システムはテーブル上の Mawari たちによる会話の場を形成し、ユーザはシステムにより展開される会話へ参加することにより、Web 上の情報へ間接的にアクセスすることができる。

3.2.1 ハードウェア

Mawari は PC の周辺機器である USB ガジェットとして設計・実装を行った。USB ガジェットであるためバッテリーレスで動作する。本体はマイコン (H8) 周辺回路と、加速度センサ、RC サーボモータと DC モータから構成される (図 2)。

Mawari の外装は、生き物らしい柔らかい動きを狙い、ウレタンゲル素材で製作した。ミニマルデザインの戦略に基づき、外装に顔や手足を表す外観はない。しかし、頷きや否定、および身体配置や姿勢の調整などの最小限の社会性の表示機構を備える。

また、センサとして三次元加速度センサが実装されており、Mawari の振動や傾きの状態を検出することで、音声以外のユーザからの行為 (叩かれたことや持ち上げられたことなど) を検出する。

3.2.2 ソフトウェア

ソフトウェアは、会話制御モジュールと振舞制御モジュールの 2 つから構成される (図 3)。

会話制御モジュールは、ユーザからの発話の認識、会話内容の音声化、および Web サービスと RSS を利用して情報の取得を行い、会話内容の生成を行う。ユーザからの発話の音声認識を行うためにオープンソースの音声認識エンジンである Julius を使用している。また、

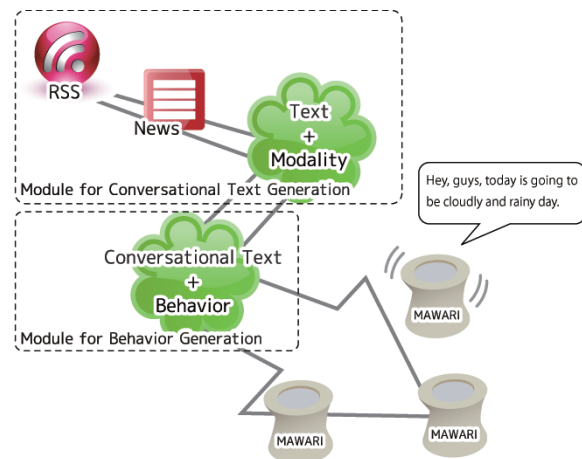


図 3: Mawari のソフトウェア概要

会話内容の音声化を行う音声合成エンジンとして ATR-Promotions の Wizard Voice を利用している。Web 上の情報の取得には Web サービスと RSS を利用しており、XML 形式で各種情報を取得することが可能となる。取得した情報を短いフレーズに分け、そこじえモダリティを付与することによって発話内容の生成を行う。モダリティとは、命題に対する話し手の心的な態度や判断を示す情報であり、命題を含む文章を話し言葉のように変換するために用いている。

振舞制御モジュールは、発話内容に基づいて各 Mawari の振る舞い動作の制御を行う。音声合成エンジンにより会話内容をスピーカから出力するとともに、モータの制御情報を Mawari へ送信する。Mawari は、PC から受け取った制御情報に基づき振る舞いの動作を行うとともに、測定したセンサデータを PC へと送信する。

4 評価実験

本システムは、会話への参加の自由度の存在から、一対一の音声インタフェースに比べて、ユーザの心的負荷が小さく、会話への参与しやすさなどが感じやすくなると考えられる。

そこで、Mawari が単独の環境を「条件 A」、複数 (3 つ) の環境を「条件 B」とし、ユーザの心的負荷 (メンタルワークロード) と Mawari への印象についての質問紙評価によって調査した。

4.1 評価方法

ユーザがシステムを利用した際の心的負荷の評価方法として、NASA-TLX (NASA Task Load Index) を用い

表 1: 心的負荷の評価項目

質問番号	尺度	説明文
Q1	知覚的要求	ニュースの内容を覚えていますか？
Q2	身体的要求	ロボット (たち) の話を聞くのに疲れましたか？
Q3	タイムプレッシャー	ロボット (たち) の話を聞くのにあせりましたか？
Q4	作業成績	ニュースの内容について理解することができましたか？
Q5	努力	ロボット (たち) の話を熱心に聞こうとしましたか？
Q6	フラストレーション	ニュースの内容がわからずにイライラしましたか？

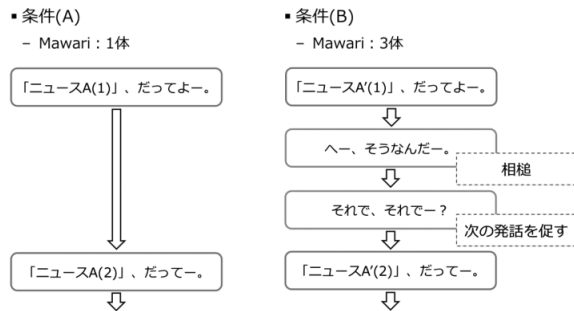


図 4: 各条件における Mawari の振る舞い

た二重課題法を設定した [9]. 評価方法としては, NASA-TLX の各尺度に加えて, 総合値として RTLX(単純平均) と AWWL(評価値が小さい順番に 1 から 6 の重み付けを行った平均) を評価に用いる.

本実験では, NASA-TLX に基づいた評価指標として, 表 1 に示すような 6 つの項目について 10 段階で評価を行うように設定した.

加えて, Mawari に対する印象の評価として, 同様に質問紙による主観評価を行う. アンケートの項目は, 親しみやすさ, 会話をしてみたいかどうか, 会話が受け入れやすいかどうかについて問うものであり, 「Q7 ロボット (たち) と会話をしてみたいと思いましたか?」「Q8 ロボット (たち) の会話の内容はあなたにとって受け入れやすいものでしたか?」について, 5 段階でユーザが評価する.

二重課題法のタスクを設定する上で, ユーザが机で PC などの作業をしながら Mawari を利用する場面を想定し, 検証課題としてニュース聞き取りタスク (Mawari が話す内容 (ニュース) を聞き, 内容について理解してもらうもの), 負荷課題として迷路ゲーム (18 × 18 マスの迷路を入り口から出口までユーザにペンでなぞってもらうもの) を設定した.

実験における会話内容は関連のない 6 つのニュースソースを選出し, こと前に生成した文を用いて, 条件 A と条件 B にそれぞれ 3 つずつ振り分けた.

条件 A では, ニュースソースにモダリティを付与した



図 5: 実験の様子 (上: A 条件, 下: B 条件)

文を連続的にユーザに提示した (図 4). 条件 B では, 条件 A の場合に加えて, 相槌や次の会話を促す文をユーザに提示した. また, 両条件の全体的な所要時間を揃えるために, 条件 A では文の間に待ち時間を設定し, 提示タイミングを調整した. 条件 A, 条件 B ともにニュースソースにモダリティを加えた文の総量はほぼ等しい. 加えて, モダリティは条件 A, 条件 B ともに同じものを順番に用いて会話生成を行った.

4.2 被験者

実験の被験者は, 男性 13 名, 女性 3 名の計 16 名 (19 歳から 21 歳, 平均 20.2 歳) であり, 全員 Mawari とインタラクションを行うことは初めてであった. 実験の実施前に教示として,

- 机の上でロボット (たち) がニュースについて話しています

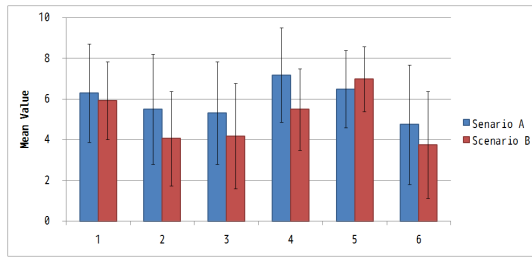


図 6: NASA-TLX 評価尺度別の結果

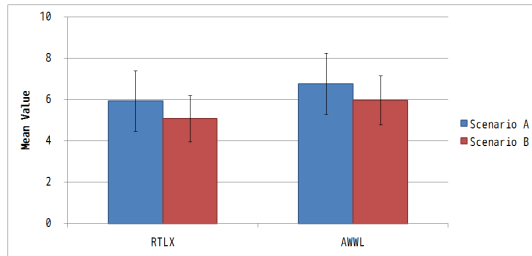


図 7: NASA-TLX 総合値 (RTLX, AWWL) の結果

- ロボット (たち) の話を聞いて、ニュースの内容について理解してください
- 同時に紙に書かれた迷路を入口から出口までペンでなぞってください

という内容を伝えた。Mawari × 1 の対話環境、Mawari × 3 の多人数会話環境の各条件についてカウンターバランスをとり、実施した。実験中の様子を図 5 に示す。

4.3 結果

評価実験の結果を図 6, 7 にそれぞれ示す。また、各平均値に対して t 検定を行った。

まず、NASA-TLX 評価尺度別の結果 (図 6) について述べる。項目 2 (身体的要求) と項目 4 (作業成績) に関して条件 B の平均値が低く、有意な差がみられた (2: $t(15)=1.88, p<.05$, 4: $t(15)=2.39, p<.05$)。他項目については有意な差はなかったものの、条件 B の場合の方が心的負荷が全体的に低くなる傾向がみられた。

次に、NASA-TLX 総合値の結果 (図 7) について述べる。RTLX と AWWL に関しては、条件 B の平均値が低く、有意な差がみられた (RTLX: $t(15)=1.84, p<.05$, AWWL: $t(15)=1.90, p<.05$)。

次に、印象評価の結果については、Q7 に関して、条件 B の平均値が高く、有意な差がみられた (Q8: $t(15)=3.50, p<.05$)。Q8 に関しては、条件 B の平均値が高い傾向があるものの、有意な差はみられなかった (Q9: $t(15)=1.65, p<.06, n.s.$)。

実験結果より、多人数会話環境である条件 B の場合に、心的負荷が低くなること、会話をしてみたいとユーザに感じさせる傾向にあることがいえる。

5 考察

いくつかの項目について有意な差が見られなかったのは、実験タスクがニュース自体を覚えなくとも進行するものであったこと (Q1)、音声合成によって生成された発話スピードは比較的ゆっくりであったこと (Q3, Q6)、などが原因として考えられる。しかし、条件 B の方が条件 A に比べ心的負荷が全体的に低くなる傾向にあると言える。

心的負荷が全体的に低くなる理由として、多人数会話の特徴である「緩やかな共同性」によってユーザの応答責任が緩和されたことが考えられる。会話はシステムにより自律的に展開されるため、ユーザは作業をしながら会話を聞き流すことができる。会話への参加には自由度が存在し、ユーザの思考を拘束することがない。そのため、ユーザの心的負荷を軽減することができたと考えられる。

条件 A においては、1 対 1 であるため、常にユーザが聞き手であることが強制されてしまう。例えばユーザは聞き手を立てていなくても、ロボットは話しかけ、ユーザがロボットに対して発話したとしても、ロボットが聞き手にならない限り会話は成り立たない。こうした、発話内容が誰にもグラウンドされない状態がフラストレーションを高めたと考えられる。一方で、条件 B においては、ユーザの役割が傍観者 (bystander) もしくは盗み聞き者 (overhearer) であることが保証される。システム側で会話がある程度閉じているために、人工的な会話内容の不自然さがあっても、「聞き流す」という選択肢があったために、ユーザの心的負荷が軽減され、また、同時に会話に加わってみたいと思わせたと考えられる。

6 応用領域

本稿で紹介する多人数会話の場に基づくソーシャルインタフェースは、次のような分野に応用可能であると考えている。

(1) 実世界指向の Web ブラウザとして:

本稿で提案したシステムは実世界指向の Web ブラウザと捉えることができる。多人数会話に基づくソーシャルインタフェースでは、我々が普段行っている会話などのコミュニケーションスキルを利用するため、コンピュータの操作方法などに関する知識を必要としない。PC のキーボードやマウスの操作ができなくとも、会話を通じてシステムを利用することができる。このよ

うな特徴を生かし、子どもや高齢者などの情報弱者層へ向けたユニバーサルインタフェースへの応用が期待される。

(2) 広告メディアとして:

ロボットを広告メディアとして応用する可能性も考えられる。近年、紙広告へ Web サイトのアドレスを載せるなどのクロスメディア型の広告や、商品やコンテンツに対するユーザのレビューを口コミ広告が一般的になっている。例えば、価格.com²というサイトでは、ユーザのレビューが商品ごとに RSS で配信されている。本研究のシステムは、Web 上の情報を RSS を介して取得することがひとつの特徴である。そのため、システムが生成した会話から直接的に情報源の Web サイトに誘導できる。もちろん、ユーザの嗜好に応じた情報フィルタリングやデータマイニングを行うことで、より効果的な広告戦略も考えられるだろう。

(3) ソーシャルメディアータとして:

初対面の他者と関わる際においては、話しかけ難い、どのように関わっていいかわからないなど、コミュニケーションの障壁が少なからず存在する。そういった障壁を乗り越える方法として、周囲の人が手本を見せるような足場作りや、共通の事柄や物を媒介することなどが考えられる。

ロボット同士がつくる多人数会話の場を、この媒介物の概念として拡張すれば、ソーシャルメディアータとしてのロボットを考えることができる [10]。

(4) ドライビングエージェントとして:

さらに、雑談によるナビゲーションや周辺情報の提供を行うことができるため、カーナビゲーションへの応用が考えられる [11]。

エージェントからの情報提供に対し、ユーザは会話による応答に加えて、ペダルやステアリングの操作により応答を返すことも考えられ、車室内においてはユーザの着座位置が固定であるため、安定したインタラクション環境を利用したシステムを設計することが可能である。

7 むすび

本研究では、雑談を介して Web 上の情報をユーザに提供するソーシャルインタフェース Mawari の提案を行った。評価実験の結果、本システムを利用した場合、心的負荷が全体的に低くなる傾向がみられた。これは私たちに親和性のある、共感的なインタフェースの実現への第一歩である。

謝辞

本研究の一部は、日本学術振興会特別研究員奨励費、科研費補助金 (基盤研究 B 21300083) によって行われている。ここに記して感謝の意を表する。

参考文献

- [1] 竹内勇剛, 岡田美智男, 角康之, 鈴木紀子: ソーシャルインタフェースという考え方, ヒューマンインタフェース学会論文誌, Vol. 7, No. 1, pp. 1-10 (2005)
- [2] 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 DII, Vol. J84-D-II, No. 6, pp. 898-908 (2001)
- [3] Mutlu B., Shiwa T., Kanda T., Ishiguro H., Hagita N.: Footing in human-robot conversations: How robots might shape participant roles using gaze cues, *Proceedings of 4th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 61-68 (2009)
- [4] 武田信也, 中野有紀子: Wizard-of-Oz による情報提供エージェントとの多人数対話における言語・非言語行動の分析, *The 23rd Annual Conference of the Japanese Society for Artificial Intelligence*, 1A2-1 (2009)
- [5] Sakamoto D., Hayashi K., Kanda T., Shiomi M., Koizumi S., Ishiguro H., Ogasawara T., Hagita N.: Humanoid Robots as a Broadcasting Communication Medium in Open Public Spaces, *International Journal of Social Robotics*, Vol. 1, No. 2, pp. 157-169 (2009)
- [6] Suzuki N., Takeuchi Y., Ishii K., Okada M.: Talking Eye: Autonomous creatures for augmented chatting, *Robotics and Autonomous Systems*, Vol. 31, No. 14, pp. 171-184 (2000)
- [7] 安田淳志, 倉本到, 辻野嘉宏, 水口充: 互いに対立する複数エージェントとの会話によるユーザの意思決定支援手法, ヒューマンインタフェースシンポジウム 2010 論文集, pp. 127-134, (2010)
- [8] 小嶋宏幸, 岡田美智男: Mawari: 多人数会話の「緩やかさ」を利用したソーシャルインタフェース, ヒューマンインタフェースシンポジウム 2009 資料集, pp. 271-274 (2009)
- [9] 三宅晋司, 神代雅晴: メンタルワークロードの主観的評価方法-NASA-TLX と SWAT の紹介および簡便法の提案-, 人間工学, Vol. 29, No. 6, pp. 399-408 (1993)
- [10] 福井隆, 竹井英行, 竹田泰隆, 岡田 美智男: GamagoriM-ikanRobot: 人と人とのつながりを引き出すソーシャルメディアータを目指して, ヒューマンインタフェースシンポジウム 2009 論文集, pp. 601-604 (2009)
- [11] 宇都裕紀, 遠藤高史, P. Ravindra De Silva, 岡田美智男, 大林真人, 塚原裕史, 吉澤 顕: 多人数インタラクションに基づくドライビングエージェントの構築, ヒューマンインタフェースシンポジウム 2010 論文集, pp. 657-660 (2010)

²<http://kakaku.com/>