

人工家族 ～ スマートホームからスイートホームへ ～

Artificial Family : From Smart home to Sweet home

杵山 祐貴^{1*} 石井 誠也¹ 佐藤 理史¹ 小川 浩平¹
 Yuki Iriyama¹ Masaya Ishii¹ Satoshi Sato¹ Kohei Ogawa¹

¹ 名古屋大学大学院 工学研究科 情報・通信工学専攻
¹ Graduate School of Engineering Nagoya University

Abstract: Recently, the low cost of devices and the improvement of the environment have made it easy for anyone to introduce smart home appliances and operate them remotely. On the other hand, the lack of widespread adoption of smart home appliances can be attributed in part to the lack of familiarity with voice commands and remote control to operate home appliances. In this research, we suggest an interface to anthropomorphize home appliances using a messaging application. We believe that by anthropomorphizing home appliances and having conversations with them as if they were family, the operation of smart home appliances will become more familiar and easier for people to use. We also devised an interpreter specialized for controlling home appliances, and built a system to interact with and operate smart home appliances via message applications.

1 はじめに

近年、半導体技術の進歩や通信規格の整備により IoT デバイスの開発が進み、スマート家電関連の技術が発展している。さまざまな研究やサービスが提案されており、特に自然言語によるコマンドを応用する技術には注目が集まっている [1] [2]。リモコンや本体に実装されたスイッチをわざわざ用いることなく、発話するだけで操作ができるため、インターフェースとして優れていると考えられているためである。そのため、家電自体と音声で対話するものや [3], Siri, Alexa などのメディアセンターとして機能するものが提案されており次第に日常的なものとなりつつある。

しかし、同時にそれらの家電が従来のスイッチやリモコンなどといった操作インターフェースを完全に代替するには至っていない。その理由として、自然言語理解の精度の問題、および、一般的に対話の対象と認識されない家電に話しかける行為の不自然さが考えられる。

まず、自然言語理解の精度の問題について、これは例えば音声認識が正確に行われたとしても発話の意味を完全に理解することが困難である、という問題である。この問題は、言語で入力する以上無数の言い回しやス

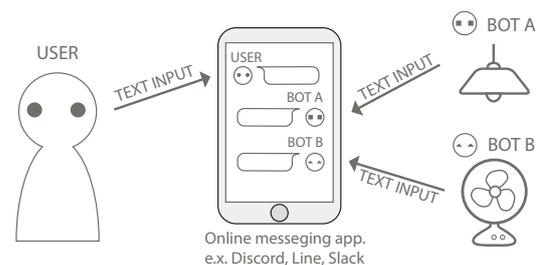


図 1: システムの概要図: 各家電はメッセージアプリ上のボットと一対一で紐づけられる。ユーザーはメッセージアプリ上の自然言語対話を通じて家電の操作が可能 (要件 1)。また、ユーザーと各家電に割り当てられたボットはそれぞれ相互にグループ対話が可能 (要件 2)。メッセージアプリは普段人間同士の対話にも使う既存のアプリケーションを使用する (要件 3)。

ラングが含まれるという問題に起因する。例えば、エアコンの温度を上げて欲しい場合、標準的には「エアコンの温度を上げて」と発話するが、場合によっては「寒いな」や「もうちょい温度上がる?」といったような発話を用いるケースもある。これらの入力を理解するためには、普段人が人同士の会話で用いるような自然言語理解が必須となり、技術的な困難性から快適な使用感という観点においてはリモコンを用いた入力が依然優位であると考えられる。

次に、対話の対象と認識されない家電に話しかける行為の不自然さについて、この問題は家電の印象に由

*連絡先: 名古屋大学大学院工学研究科
 情報・通信工学専攻 佐藤・小川研究室
 〒 464-0815 愛知県名古屋市千種区幸川町 2 丁目 8-2 ライオンズマンション名大ウエスト 408 号室
 E-mail: iriyama.yuki@d.mbox.nagoya-u.ac.jp

来する問題である.. 一般的に人は人工物の持つ見かけから、人工物のもつ知性や対話能力を推定し、それに従った方法によって関与を試みる [4]. そのため、通常対話の対象と見なされない、冷蔵庫、炊飯器、および照明器具などの家電に対して言葉で語りかけることに対して、一定の躊躇が生じる可能性がある。また、逆に家電から言葉により語りかけられることについても、違和感を与える可能性がある。

このように、家電の中心的なインタフェースとして言葉を用いることが受け入れられるためには、未だ多くの問題の解決を待たなければならない。そこで本研究では、家電を対話の対象として認知し、一定の自律性をもったエージェントであるとユーザが見なすことができるシステムの実現を通じて、これらの問題の解決を目指す。

これまで、家電などの非生物をエージェント化する試みは提案されている。例えば大澤らは、冷蔵庫にロボットの目と手をつけることで家電を物理的な方法でエージェント化している [5]。しかし、日常、数多く用いている家電全てにこのような仕組みを入れることは、現実的には困難である。また、小川らは、メディア間を移動可能なエージェントをもちいて、家電のエージェント化を試みている [6][7]。しかし、発話や明滅といった振る舞いを家電に付与する必要があるため、大澤の研究同様、既存の家電に応用することは困難である。

そこで本研究では、Line, Slack, Discord などの既存のオンラインコミュニケーションツール上で、家電とユーザが対話することで、一般的な家電に手を改めて加えることなく、家電を対話の対象として認知することができる、家電のエージェント化システムを実現する (図 1)。これにより、家電を対話の対象とみなし、人同士と似通った方法で家電と人がかかわることができる、新しい人と家電の関係性の構築が可能であると考える。ここで、既存のオンラインコミュニケーションツールを用いる理由は、日常的に用いられるからこそ、テキストでのコミュニケーションであっても人との関わり合いと似通った現実性を感じることができる可能性があるからである。

本稿では、2章において目的の達成に向けて作成したプロトタイプシステムの概要とエージェントの対話戦略について説明する。また、3章において現状でのシステムの具体的な動作例について紹介し、今後の実装方針について述べる。

2 家電のエージェント化システム

2.1 システムの概要

本研究の目的は、家電を対話の対象として認知させることで、人同士と似通った方法で家電とのかかわり

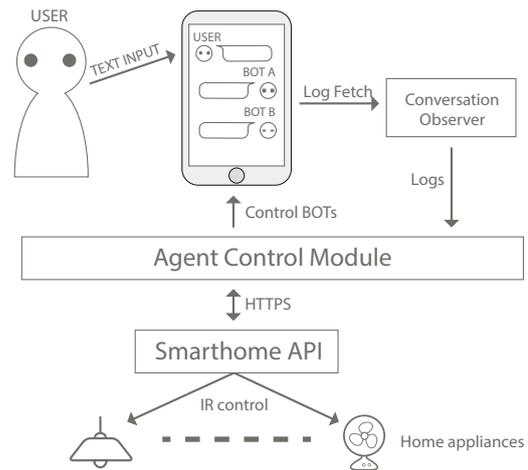


図 2: システムの構成図: 中央の“Agent Control Module”が API を介してメッセージアプリ上の対話ログを取得し、HTTP 通信によりメッセージアプリ上の Bot と家電制御用のスマートリモコンをコントロールする。

合いをもつことができる、「家電のエージェント化システム」の実現である。そのため、本稿では、図 1 に示すように、ユーザと家電が、メッセージアプリケーション上での対話を通じて、ユーザから家電への依頼、家電からユーザへの依頼、ユーザと家電との雑談、家電間での雑談を通じた情報伝達などの機能の実現に向けたプロトタイプシステムの実現を目指す。

具体的には次の 3 つの要件を満たすことで、家電のエージェント化を目指す。

- 要件 1** 自然言語による対話を通じて、家電を現実に動作させることができる
- 要件 2** 雑談対話を通じて、家電との関係構築ができる
- 要件 3** 既存のメッセージアプリケーションを対話のプラットフォームとして用いることで、対話の対象である家電に人らしさを見出すことができる

「1. 対話を通じて、家電を動作させることができる」について、提案するインタープリタによるユーザー入力理解の実現と API を用いたスマートリモコンの遠隔制御により実現した (2.2 節)。「2. 雑談など家電の操作に限らずさまざまな対話ができる」について、複数の家電ボットとのグループ対話を分類した (2.3 節)。「3. ユーザーが人間とコミュニケーションするときと同じ方法で家電と対話できる」について、技術検討とボットの実装を行った (2.4 節)。

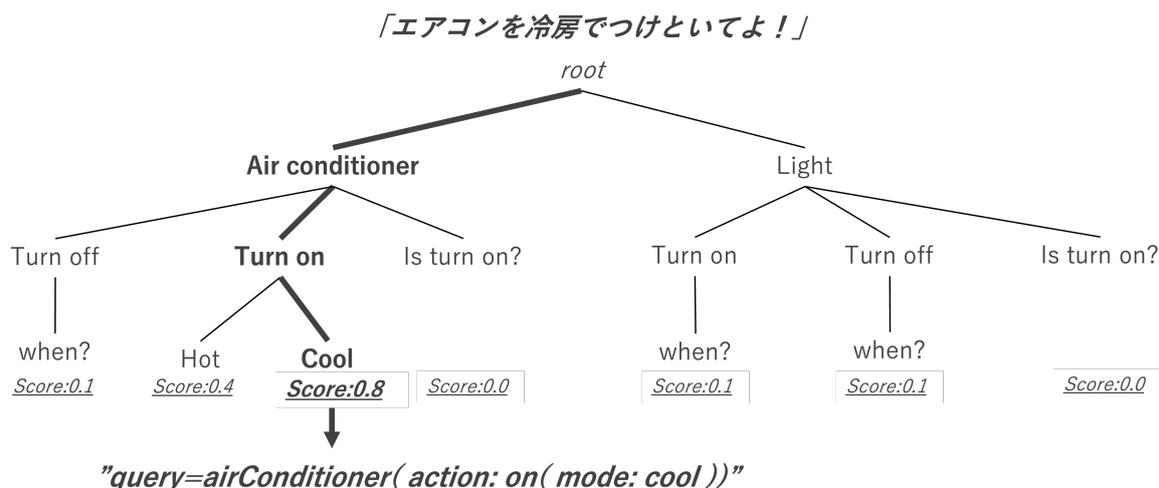


図 3: エージェント制御モジュールのインタープリタ: 自然言語による入力を, 家電操作コマンドに翻訳する. 例えば, “air conditioner” は “Turn off”, “Turn on”, “Is turn on?” の 3 種類の操作を受け付けており, “Turn on” は “Hot” と “Cool” の属性を受け付けている, ということを表している.

2.2 システム構成

システム構成を図 2 に示す. 本システムの動作について, Bot 管理, インタープリタ, 家電制御の点から説明する.

2.2.1 Bot 管理

Bot の振る舞いを管理するために, 既存のメッセージアプリケーションである Discord 上での, ユーザとエージェントの対話履歴を取得し, “Agent Control Module” に提供する. これにより, 先述の「要件 3」で述べた人との関わり合いと似通った現実性をエージェントとの対話に感じさせることができると考える.

具体的には, 本モジュールはメッセージアプリ上の Bot の制御とチャットスペースの監視・トークログの回収を行う. Bot の制御は webhook によるメッセージの送信機能を活用し実現した. チャットスペースの監視・トークログの回収は Discord.py を用い, ユーザーの発言をトリガーに発言内容・時間・発言者などの情報をまとめて “Agent Control Module” に送る実装を行うことで実現する.

2.2.2 インタープリタ

インタープリタは, 図 2 中に示す “Agent Control Module” 内に実装されており, “Conversation Observer” から送られてきた会話ログを受け取り, Bot の返答生成と家電操作コマンドの生成を行う. ユーザー入力の解釈は MeCab/CaboCha による係り受け解析結果をもとに

「対象の家電 (エアコンや電灯など)」「操作の種類 (電源をつける, 温度を下げるなど)」「付随する属性情報 (タイマー設定時間, 設定温度など)」を含む家電制御モジュールが理解できる Json 形式の家電操作コマンドを生成する. 詳細なインタープリタの構造を以下に示す. インタープリタはシステムに組み込まれている家電の種類, 操作, 属性を木構造で保持している. (図 3) たとえば “air conditioner” は “Turn off”, “Turn on”, “Is turn on?” の 3 種類の操作を受け付けており, “Turn on” は “Hot” と “Cool” の属性を受け付けている, ということである.

次に, インタープリタの処理のアルゴリズムについて説明する. この木構造の各ノードは入力文に対して「自分について言及しているかどうか」を示すスコアを算出する. 例えば入力文の動詞に反応して Turn on は高いスコアを算出し, 同様にして Turn off は低いスコアを算出する. 次に, 各段で最も高いスコアを出したノードを選択しながらルートから葉ノードまでのパスを選択する. 最終的に, 選択されたパスに従って各ノードからコマンドを生成する.

この構造は新しい家電の追加に対して強く, またスコアを動的に変更するアルゴリズムを導入すれば状態を更新しながらコマンドの生成が可能となる拡張性を備えている.

また, 決定木を用いた解釈ではなく, 得点を意味解釈の蓋然性ととらえることで, 入力の意味を一意に決定せず複数の可能性を残した柔軟な解釈が可能になるため, 自然言語に含まれる曖昧さのある程度許容することができる.

2.2.3 家電制御

家電の制御は、スマートホームを遠隔操作可能にする API である SwitchBotAPI に対して HTTPS リクエストを飛ばすことで、部屋に設置されたスマートリモコンデバイスを操作し、実世界の家電に赤外線信号を伝達し動作させることで実現する。

具体的には、“Agent Control Module” の生成したコマンドから、対象の家電や温度の情報などを抽出し SwitchBotAPI 用のクエリを生成する。また、タイマー設定など SwitchBotAPI で制御できないパラメータについてはシステム内でリクエストのタイミングをコントロールすることで制御している。

現在システムはエアコン、電灯の制御に対応している。

2.3 ユーザーとエージェントの対話戦略

要件 1 と要件 2 を実現するエージェントの対話戦略を構築するために、ユーザーとエージェント複数から構成されるおけるグループ対話に関する分析を行った。具体的に、「対話の形態」と「対話の対象」という二つの視点から分類し 6 パターンの対話モデルを構築した。ここで、本分類における会話とは、1~3 ターンほどの短期的な会話を想定する。

2.3.1 タスク指向型会話と雑談型会話による分類

対話をタスク指向型の会話と雑談型の会話に分類することで、コマンドによる家電制御や情報の通知など明確なタスクを目的とするか、エージェントとユーザーの関係構築を目的とするかを明確にする。具体的には、「エアコンをつけて頂戴」や「明日ゴミ出しの日だよ」など目的の存在する会話をタスク指向型会話、「ただいま」や「留守にしている間どうしてた？」など目的の存在しない会話を雑談型会話と分類する。2 章で説明したメッセージアプリ上での会話による家電操作はタスク指向型会話に分類される。

2.3.2 対話の話し手と聞き手による分類

「誰が誰に対話を行うか」という視点から三種類の分類を行った。この分類により、ユーザーと複数エージェント間で行われるグループ対話で発生する対話パターンを網羅的に把握し、前節の分類と合わせて家電のエージェント化という目的においてどのような役割を持たせられるか分類した。

Human-Agent 型会話

Human-Agent 型会話 (以下 H-A 型会話) は人間がエージェントに話しかける構造の会話である。タスク指向型会話の H-A 型会話の例としては 2 章で説明したメッセージアプリ上での会話による家電操作が挙げられる。雑談型の H-A 型会話は「ただいま」のようなエージェントに対する明確な期待のない会話や、「なんか面白い話して」のようなエージェントとの関係構築を目指す会話が例として挙げられる。

Agent-Human 型会話

Agent-Human 型会話 (以下 A-H 型会話) は対称的にエージェントが人間に話しかける構造の会話である。例えば、エアコンの長時間稼働を検出して「エアコンつけすぎで電気代心配だよ」のような通知を行うタスク指向型会話が考えられる。また、「なんか好きな音楽教えてや」や、「おなかすいた〜」のような雑談型 A-H 型会話を適切に取り入れることが家電のエージェント化を促進する要素となる可能性も検討している。

Agent-Agent 型会話

Agent-Agent 型会話 (以下 A-A 型会話) は参加者がエージェントのみに限定される会話である。ユーザーは直接会話に参加することはないが、会話の流れを見て情報を得ることは可能である。A-A 型会話は、会話の内容を全てシステムが事前に決定できるため破綻の少ない会話を展開できるというメリットが存在するほか、一部の情報伝達に際して A-H 型会話と比較してより効果的であるとされている [8]。具体的な活用方法としては、「明日雨降るんだって」のようなユーザーへの情報提示 (タスク試行型) や「面白い記事があってね」のようなエージェント化を促進する雑談型会話を想定している。

2.4 エージェントとの対話媒体

家電のエージェント化に際して、ユーザーが他の人間とコミュニケーションするときと同じ媒体を用いることは有用であると考えられる。

対話媒体の選定において「テキストコミュニケーションを行う媒体であること」「複数のエージェントによるグループ対話が可能であること」「日常的に家族や友達の間で使う媒体であること」が求められた。上記の条件を満たす媒体として、本研究では Discord をメッセージアプリケーションに採用した。Discord は音声通話機能とテキストチャット機能を組み合わせたゲーマー向

	入力文	インタープリタ出力の解釈	返答文	家電の動作
1	電気をつけて	対象: light 操作: turn on 属性:	はい	部屋の電気が点灯した.
2	電気消しといて頂戴!	対象: light 操作: turn off 属性: N/A	はいはい	部屋の電気が消灯した.
3	22時になったらエアコン, 暖房 21度で焚いといて頂戴~	対象: air conditioner 操作: turn on 属性: 暖房, 22:00, 21度	おっけ	設定通りにエアコンが 起動した.
4	電気とエアコンつけといて頂戴	対象: light 操作: turn on 属性: N/A	はいよー	部屋の電気が点灯した.

表 1: システムの入力文と生成された解釈および返答文: 1~3 は想定した解釈が得られた. 4 については対象となる家電の認識に失敗している. 返答文は既定のリストからランダムに出力される.

けコミュニケーションツールであり, 2体以上の Bot を一つのチャットグループに導入して会話することができる.

3 システムの動作例と今後の実装方針

3.1 ユーザーの入力の解釈例

システムの動作例を表 1 に示す, 入力文は Discord 上に作成した専用サーバー (チャットグループ) に対して入力された. インタープリタの解釈はコマンドラインに出力した json 形式から内容を抽出したものである. 返答文は上記の専用サーバー上から回収された.

表 1 中 1~3 について, 入力に対して意図したとおりの出力が得られた. 表 1 中 4 については一つの入力に対して複数のインタープリタ出力を得ることを目的とした入力である. 現状インタープリタは一つの入力に対して一つの出力しか生成しないため, 正しい出力を得られていない.

3.2 今後の実装方針

3.2.1 エージェント同士の会話によるユーザーの行動変容

グループ内で 2 者の対話が行われる状況において, 対話の聞き手に回るケースと比較して対話の第三者として聞くケースの方が対話の内容が効果的に伝わるとする先行研究が存在する [8]. 本研究においても, 通知内容のような家電からユーザーに情報を伝達する際において Bot 同士の対話による情報伝達が有効であると考

える. A-A 型対話 (2.3.2 節) において, 通知などのトピックに対して 2 社対話を構成する技術を用いることでスクリプトを作成し, メッセージグループ内のボット同士を対話させることでトピックの伝達を目的としたタスク指向 A-A 型対話を実現する.

3.2.2 エージェントとユーザーの非構造会話によるエージェントの印象変化

家電のエージェント化において, タスク指向型の会話だけでなく雑談型の対話も行えるシステムを開発することは, ユーザーの家電に対する認識を変化させるうえで重要であると考えている. また, 雑談型の対話における返答やテーマ設定に偏りを持たせることでエージェントにキャラクター性を付与することができる可能性がある. 以上を検証するために, 雑談型の対話に対応したモジュールを開発するがある. 詳細には, H-A 雑談型の対話に対応するための雑談チャットボットモジュール, A-H 雑談型対話と A-A 雑談型対話を生成するための雑談トピック生成モジュールの制作が必要である. 現在, 前者の H-A 雑談型対話返答モジュールは杉山らの開発した機械学習によるチャットボット [9] を活用して開発を進めている.

3.2.3 実験手法と評価指標の検討

仮説検証のために, 開発したシステムを用いた事件とその評価が必要である. 本システムの実験評価は, 1. 多数の被験者を集めた実験が難しい点, 2. 定量的な評価が難しい点において問題が存在する. 1. 多数の被験者を集めた実験が難しい点について, 本研究は長期的なシステムの使用において家電のエージェント化が発

生するかを検証する必要がある。多数の被験者に対して長時間追跡調査することは困難である。また、本システムは生活で使用する家電設備上で駆動するため被験者毎に技術的に適切な設定を行う必要がある。多数の被験者に対してそれぞれ家に赴いて機器のセッティングを行うことは現実的ではない。

2. 定量的な評価について、エージェント化は主観的な現象のため客観的指標が存在しない。また、主観評価においても上記の問題により十分なサンプル数をあつめることが困難である。これらにより、実験結果に基づく定量的な比較・評価は困難である。

現時点では、エスノメソドロジーの手法を用いて少数の被験者によるシステムの評価を行うことを検討している。

4 結論

本論文では家電のエージェント化に必要な要素を列挙・考察した。また、上記の考察に基づき自然言語を用いてチャットアプリのボットに話しかけることで家電が操作できるシステムを構築した。

今後はユーザーとのボットを通じたコミュニケーション経験をの向上を目指し、家電のエージェント化を促進する。具体的には2.3で整理した対話の分類に基づき、ユーザーと家電エージェントがより多様な対話を行えることを目指し、機能の追加を行う。

また、平行してシステムを通じた家電のエージェント化がどの程度実現されるかについて実証実験を行い分析・フィードバックを得ることも予定している。また、実証実験に伴いシステム評価手法の提案と実験手法の立案が必要である。

最終的に本研究を通じて家電のエージェント化を実現し、1人暮らしのマンションにおいても家族のような家電エージェントに囲まれた「スマートホーム」を経験できるようなシステムの構築手法の提案を目指す。

参考文献

- [1] 西山高史. スマートホーム技術の現状と将来—SoS視点も入れて—. 計測と制御, Vol. 59, No. 12, pp. 957–960, 2020.
- [2] 総務省 | 令和元年版 情報通信白書 | IoT デバイスの急速な普及.
- [3] 小川耕司, 山下浩司, 小暮悟, 小西達裕, 伊東幸宏. メニュー操作を考慮した機種非依存の家電統合インタフェース. 情報処理学会研究報告音声言語情報処理 (SLP), Vol. 2008, No. 12(2008-SLP-070), pp. 89–94, February 2008.
- [4] 星田雅弘, 田村昌彦, 林勇吾. 身体性を有するエージェントとのコミュニケーションにおける言語選択. 電子情報通信学会技術研究報告; 信学技報, Vol. 116, No. 436, pp. 71–76, January 2017. Publisher: 電子情報通信学会.
- [5] 大澤博隆, 向井淳, 今井倫太. 身体イメージを想起させる擬人化ディスプレイロボット. 情報処理学会研究報告. ICS, [知能と複雑系], Vol. 146, pp. 1–6, December 2006. Publisher: 一般社団法人情報処理学会.
- [6] 小川浩平. ITACO : メディア間を移動可能なエージェントによる遍在知の実現. ヒューマンインタフェース学会論文誌, Vol. 8, No. 3, pp. 373–380, 2006.
- [7] ITACO on the Room: アンビエントな情報提供を行う生物感のあるエージェントの提案. p. 6.
- [8] 磯和隆道, 小川浩平, 佐藤理史, 石黒浩. 情報の対話的再構成システムの実現. 人工知能学会全国大会論文集, Vol. JSAI2020, pp. 2P4GS1105–2P4GS1105, 2020.
- [9] Hiroaki Sugiyama, Masahiro Mizukami, Tsunehiro Arimoto, Hiromi Narimatsu, Yuya Chiba, Hideharu Nakajima, and Toyomi Meguro. Empirical analysis of training strategies of transformer-based japanese chat systems, 2021.