

自律移動型ロボットにおける 経験由来の情報量からの発話内容選択

吉原 大智^{1,2}, 湯口 彰重^{2,3}, 河野 誠也², 飯尾 尊優^{4,2}, 吉野 幸一郎^{2,3}

¹ 筑波大学 ² 理化学研究所 GRP ³ 奈良先端科学技術大学院大学 ⁴ 同志社大学

s2120792@es.tsukuba.ac.jp, tiio@mail.doshisha.ac.jp

{akishige.yuguchi, seiya.kawano, koichiro.yoshino} [at] riken.jp

概要

ユーザとの共生空間で動作するロボットには、ユーザから信頼され協調動作をスムーズに実行することが求められる。そのためには、ロボットがどのような外界観測を行いそこから何を記憶しどのような意図を生成したのかを、ユーザが理解できるように表出させることが重要である。外界観測結果や記憶・意図を表出させる有効な手段として、言語を用いたインタラクションがある。長期的なインタラクションを想定したとき、冗長なインタラクションを避けるために、何を言語化し伝えるかだけでなく、何を言語化せず伝えないかも重要である。本研究では、経験由来の情報量（エントロピー）をロボットが言語化する内容の選択に用いる。具体的には、GPT 由来の一般言語モデルとロボット自身の経験から学習した言語モデルの2種類を利用する。ロボットを使った案内タスクで被験者実験を行った結果、提案法がロボットの内部状態を必要十分に説明できることが示された。

1 はじめに

ロボットや人工知能技術の進展に伴い、人間の生活空間で自律的に動作するロボットが実現されようとしている。こうした自律型ロボットが人間の生活空間で動作しようとする場合、ロボットの状態がどのようなものか、人間のユーザにとって明確である必要がある[1, 2]。例えば、ロボットが何をでき何を記憶しているのかということをユーザが理解できなければ、ユーザは適切な命令をロボットにできず、協調作業などに支障が出る。言い換えれば、ロボットの内部状態が、ユーザにとって必要なタイミングで理解可能であることが重要である[3, 4]。

しかし、ロボットは人間とは異なる動作原理で動

作しており、人間であるユーザがロボットの意図や行動の理由を読み取ることはしばしば難しい[5, 6]。言い換えれば、ユーザにとってロボットの行動予測は困難であり、ユーザにとってはいつ何をどのようにロボットに要求してよいかわからないというユーザビリティの問題が生じる。こうした不可視な内部状態の問題は、ロボットが得られる信頼性の問題にも繋がっている[7, 8, 9, 10]。

この問題を解決する方法のひとつが、外界観測結果・記憶・意図といったロボットの内部状態をユーザに自然言語を用いて対話的に説明することである[11, 12, 13, 14, 15, 16, 17, 18]。この考え方は、Explainable AI (XAI) の考え方をロボットに当てはめたもので、ユーザとの協調性を高めるための手法として注目されている[4, 6, 13, 19, 20, 21, 22]。

しかし、ロボットの内部状態は観測結果・記憶・動作計画・バッテリー残量など種類が多く、実際にはどの情報を言語化（発話）し説明することがユーザとの協調動作に有効かは明らかではない。無分別な言語化・説明はユーザビリティの低下に繋がるため、場面に応じた適切な量・内容の言語化が必要である[22, 23]。つまり、ロボットの内部状態のうち、何を言語化する・しないを決める基準を定め、説明することが重要である。

ロボットの内部状態を説明するとき、何を言語化するか決定する基準に関する研究では、ロボットの安全性に関する情報を重視することや[24]、ロボットのタスク遷移の要点を重視することが提案されている[21]。これらの研究は、主にユーザとロボットの初対面でのインタラクションを想定している。しかしユーザの生活空間での活動するロボットは、より長期的な観測・記憶を活用する必要がある場面も多い[25, 26]。また、ロボットの行動パターンがルーティンである場合、その報告は冗長となる一方

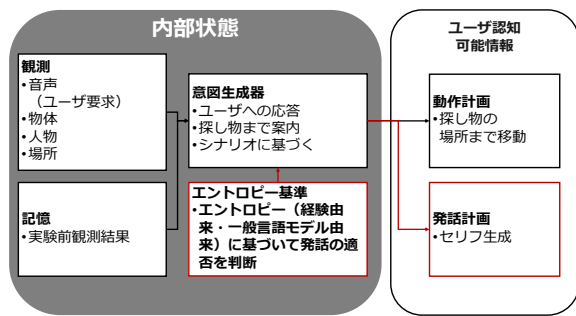


図1 ロボットの内部状態とモジュール

で、行動パターンが何らかの理由でいつもと異なる場合はそれを報告する必要がある。つまり、生成した説明の必要十分性を担保するよう、ユーザの慣れや状況に応じて過不足なく説明を行うという視点が必要である [27]。

本研究では、過去のユーザとの対話をもとにロボットが発話する対話戦略の下、発話内容がある基準によって削ることは説明の必要十分性の向上に有効なのかを検証する。ここで、この発話内容の選択に用いる基準のことを言語化基準と呼ぶ。この際、過去の記憶・経験と比較した場合の現在の観測の希少さを言語化基準として用いる。具体的には、過去の記憶・経験を利用した確率的言語モデルを構築し、この言語モデルに対する現在の観測事象のエントロピーを計測する「エントロピー基準」を用いる。このエントロピー基準によって説明の必要十分性が向上するか、ロボットを用いた被験者実験で評価・検証する。

2 タスク

本研究では、ロボットによる言語化の適切さを測るため、自律型ロボットにおいて案内タスクを設定し実験を行った。自律移動型ロボットはリビングを模した実験室内で案内タスクを行う。案内タスクとは、ユーザ役の実験者がある物体を探しており、ロボットにその物体を探していることを伝え、ロボットがその場所を教えるタスクである。ロボットは自身の観測と記憶からなる内部状態に基づいて物体がある場所を推定し、その物体がある地点まで案内する意図に基づき行動計画と発話計画を生成する。このとき、ロボットは自身の認識結果と言語化規則（詳細は付録の表2）によってあらゆる認識結果を言

語化する。この言語化内容のうち、何をユーザに対して発声し、また何を発声しないか決定するというのが本論文で取り組むタスクである。つまり図1で示すように、ロボットが持つ多様な観測結果が言語化された結果に対して、発声するかどうかを選択するフィルターを実装する。

3 情報量に基づく発話内容選択

本研究では、ロボットが生成した発話が、過去の記憶・経験と比較した場合にどの程度希少かに基づいて発話すべき内容を決定することを考える。このとき、確率的言語モデルにおけるエントロピー、あるいはパープレキシティ (*ppl*) の考え方をを用いる。本章では各言語化候補（発話候補）に対してスコアを付与する方法について述べる。

3.1 言語モデルとパープレキシティ

今回用いるロボットは現在の観測や行動に対して言語化を行う規則を有するため、あらゆる内容がテキストとして保存される。このテキスト中に含まれる単語列 s の同時生成確率を $P(s)$ とする。過去の観測結果の集合である \mathcal{S} から最尤推定、あるいはMAP推定に基づいてモデルパラメータ θ を構築する。このパラメータを用いて現在言語化を行うか考慮している文 s に対する尤度 $P(s; \theta)$ を考える。このとき、この尤度が低いということは過去の観測集合 \mathcal{S} からは遠い文、つまり希少性の高い文であると言える。逆に、この尤度が高い場合は希少性の低い文であり、ロボットからすると言語化する価値の低い可能性が高い文という仮説が成立する。

この尤度計算には計算機のまるめ誤差の問題から一般に負対数尤度を用いるため、文 s に対する尤度、つまりエントロピーおよびパープレキシティは以下の式で計算される。負の対数尤度なので数の大小は反転しており、エントロピーあるいはパープレキシティが高い方が希少性の高い文ということになる。

$$H = -\log P(s; \theta) \quad (1)$$

$$ppl = 2^H \quad (2)$$

ここで θ の計算方法について2種類を試行する。一つ目は、ロボット自身の過去の経験のテキスト集合 \mathcal{S} から最尤推定によってパラメータを推定する方法である。もう一つは、Web上に存在する大量テキストから学習された大規模言語モデル（一般言語モデル）が与えるパラメータを用いる場合である。前

者はロボット自身の経験上の希少性を用いることに相当し、後者は世の中一般の常識における希少性を利用することに相当する。

3.2 経験由来パープレキシティ

ロボットが過去に見聞きした内容の集合 \mathcal{S} に基づいて発声候補のパープレキシティを計算したものを以降では ppl.exp と記載する。 ppl.exp は、2022 年 11 月 24 日時点でぶつくさ君の行った外界観測の結果から遡って物体認識 24,278 個、人物認識 917 個、音声認識 650 個、場所認識 385 個を規則に基づいて言語化した 407,535 文から学習した。これらの文を `sentencepiece`[28] でトークナイズし、`fairseq`[29] で言語モデル化した。

3.3 一般言語モデル由来パープレキシティ

一般言語モデルに基づいて発声候補のパープレキシティを計算したものを以降では ppl.GPT と記載する。一般言語モデル由来のパープレキシティの計算には、`rinnna` 社が公開している言語モデル `rinnna/japanese-gpt2-medium`[30](以下 `lm_general` と呼称) を利用した。経験由来パープレキシティと合わせて、`huggingface`[31] により計算を行った。

3.4 パープレキシティの閾値

実際に各モデルからパープレキシティを計測したとして、発声する・しないの閾値をどう設定するかという問題がある。この閾値は、経験由来パープレキシティの計算に用いたデータ \mathcal{S} とは別の開発用データ (2022 年 11 月の物体認識 174,189 個、人物認識:15,108 個、音声認識:40 個、場所認識:66 個) を利用して計算する。ここから言語化規則に基づいて作成された 1,323,448 文に付与可能な ppl.exp と ppl.GPT をそれぞれ計算し、その平均を閾値とした。

4 実験

実際に案内タスクでロボットに発話を行わせ、その様子をユーザに見せる被験者実験を行った。本節では比較に用いた条件、評価指標、結果について述べる。

4.1 比較条件

情報量基準に基づく発声候補の選択がユーザからの印象にどのような影響を与えるかについて検証する。今回は経験由来と一般言語モデル由来の 2 種

類の手法が存在するため、以下の候補を比較条件とした。

- 条件 1: 全ての言語化を発声する
- 条件 2: 全ての言語化を発声しない
- 条件 3: ppl.exp が低いものを発声しない
- 条件 4: ppl.GPT が低いものを発声しない
- 条件 5: ppl.exp が高いものを発声しない
- 条件 6: ppl.GPT が高いものを発声しない
- 条件 7: 条件 3 かつ条件 4 のものを発声しない
- 条件 8: 条件 3 または条件 4 のものを発声しない

ここで条件 5 は条件 3 の、条件 6 は条件 4 の反転条件である。また、条件 7,8 はそれぞれ条件 3,4 を組み合わせた `and/or` 条件である。

4.2 実験条件・評価指標

評価では、ロボットが実際に案内タスクを行っている様子を記録した動画を収録し、その動画をクラウドワーカーに見せることによって点数をつけてもらった。映像は 3~4 分程度となっている。なお、映像ではユーザは花瓶を探しているものとした。

評価項目としては、動画視聴の事前事後として以下の項目を付与してもらった。

事前アンケート

- ユーザの年齢・性別
- ロボット否定的態度尺度 (14 項目 5 尺度)[32]

事後アンケート

- 発話内容 (5 項目 5 尺度、1 項目 7 尺度)
 - 説明の適切性 (1 項目 5 尺度)
 - 説明の必要十分性 (1 項目 7 尺度)
 - 主体性 (1 項目 5 尺度)
 - 会話性 (1 項目 5 尺度)
 - 説明の選択性 (1 項目 5 尺度)
 - 説明内容の珍しさ (1 項目 5 尺度)
- TAM (20 項目 5 尺度)[33]
 - 使用時の不安 (4 項目 5 尺度)
 - 利用しようとする意向 (3 項目 5 尺度)
 - ロボットの適応能力の知覚 (3 項目 5 尺度)
 - 使いやすさの知覚 (5 項目 5 尺度)
 - 有効性の知覚 (3 項目 5 尺度)
 - 技術の信頼性 (2 項目 5 尺度)
- 自由記述

5 尺度の項目は 5 が最高、1 が最低である。「説明の必要十分性」のみ 7 段階になっており、4 がもっと

表 1 実験結果。必要十分性については、目標値である 4 との差分も記載

条件	評価指標						TAM					
	適切性	必要十分性	主体性	会話性	選択性	珍しさ	不安	利用意向	適応能力	使いやすさ	有効性	信頼性
条件 1	4.35	6.05 (+2.05)	3.80	3.65	3.25	1.90	7.95	8.70	10.90	17.90	9.60	6.65
条件 2	2.20	2.40 (-1.60)	3.25	2.00	2.35	2.15	11.10	7.45	8.50	17.30	8.75	6.05
条件 3	4.15	5.10 (+1.10)	3.75	3.50	3.85	2.05	8.40	9.50	10.90	19.20	10.40	7.25
条件 4	3.45	3.75 (-0.25)	3.10	2.95	2.85	2.20	9.60	8.65	9.70	17.45	9.30	5.90
条件 5	4.35	5.75 (+1.75)	3.20	3.60	3.35	2.00	8.40	9.20	10.60	19.20	10.15	7.05
条件 6	4.05	6.00 (+2.00)	3.75	3.55	3.50	2.15	9.55	7.95	10.45	18.65	9.40	7.15
条件 7	4.10	5.20 (+1.20)	3.25	3.45	3.40	2.10	8.80	10.50	10.90	19.25	10.60	7.00
条件 8	2.15	2.25 (-1.75)	2.80	1.65	2.40	2.30	12.30	6.30	7.95	15.20	7.10	4.55

も必要十分、1 が最も説明が足りない、7 が最も説明が冗長となる場合に対応することに注意されたい。

参加者は、2022 年 12 月にクラウドソーシングで、全条件を同じ参加者に視聴してもらい比較評価するのは難しいため、各条件 20 名ずつ計 200 名 (女性 102 名、男性 98 名、19 歳～65 歳、平均 41.6 ± 9.19 歳) を集めた (条件 1: 平均 42.7 ± 11.0 歳、条件 2: 平均 39.5 ± 10.6 歳、条件 3: 平均 39.7 ± 6.8 歳、条件 4: 平均 41.2 ± 11.5 歳、条件 5: 平均 39.6 ± 8.2 歳、条件 6: 平均 41.2 ± 8.7 歳、条件 7: 平均 43.1 ± 10.0 歳、条件 8: 平均 44.9 ± 7.5 歳)。

4.3 結果

まず、事前アンケートの各項目について、各条件間でクラスカル・ウォリス検定を行ったところ差は見られなかった。実験に関わる各アンケートについての実験結果を表 1 に示す。表から、今回必要とする説明の必要十分性に関しては、発声内容のフィルタをしない場合 (条件 1) と比較して、フィルタをする場合はいずれもスコアが向上する (4 に近づく) ことが確認された。特に ppl.GPT のみを使う場合 (条件 4) が最もロボット発話の必要十分性を向上させることがわかった。またこのうち、スコアの絶対値を比較した場合には条件 1 対条件 2、条件 1 対条件 4、条件 1 対条件 8、条件 3 対条件 4 の組み合わせで、スティール・デュワス検定 ($p < 0.05$) における有意差が確認された。この結果から、全ての観測内容について言語化を行うよりも、一般言語モデルを用いて発声内容をフィルタする方が、対話ロボットの説明に感じる必要十分性が向上することがわかった。これらの場合いずれも、スコアの絶対値で比較すると有意差があるものの、4 との差分の絶対値で見た場合には有意差はない。また、提案法の逆転条件である条件 5、条件 6 のスコアは 4 から大きく離れ、提案法の有効性を示唆する結果となった。

その他の項目では、発声内容のフィルタをしない場合 (条件 1) が説明の適切性、主体性、会話性の上で高いスコアを得た。これらの項目は、ロボットの

説明に感じる必要十分性とトレードオフの関係があると考えられる。説明の選択性については ppl.exp のみを使う場合 (条件 3) が高い評価を得たが、これはロボットが自身の経験に基づいて発声内容を選択した結果、ユーザからもそうした選択をしていることが強く感じられたのではないかと考えられる。

また、各 TAM 尺度についても比較評価を行った。全体の傾向としては、and 条件 (条件 7) が高いスコアを得る結果となったが、信頼性においては ppl.exp (条件 3) が一番高いスコアを得た。発声内容のフィルタをしない場合 (条件 1) と比較すると、いくつかの項目では同等のスコアを得ている中で、それ以外の項目ではフィルタをする条件の方がスコアが向上している場合が確認できる。また、使用時の不安においては何も発声しない場合 (条件 2) や or 条件 (条件 8) が高いスコアを得ている。この結果から、必要以上に発声を抑制することはユーザにとっての不安に直結することが見てとれる。

5 おわりに

本研究は、長期的なロボットとのインタラクションを想定したときに必要となるであろう、ユーザにとって有益な内部状態のみを発話させるロボットの実現のために、発話選択を発話を持つ情報量に基づいて行う枠組みを提案した。結果として、一般言語モデル由来の情報量を用いることで、ロボットが発生する内容の必要十分性を向上させることができた。今後は、ロボット自身の経験などをどのように併用することが適当か検討する必要がある。また、特定ユーザを対象とした長期的なインタラクションを行う場合についても検討する。

謝辞

本研究は JSPS 科研費 22H04873 の支援を受けた。

参考文献

- [1] Nikhil Churamani, Sinan Kalkan, and Hatice Gunes. Continual learning for affective robotics: Why, what and how?

- In **Proc. RO-MAN**, pp. 425–431, 2020.
- [2] Alessandra Rossi et al. How social robots influence people’s trust in critical situations. In **Proc. RO-MAN**, pp. 1020–1025, 2020.
 - [3] Xiaofeng Gao et al. Joint mind modeling for explanation generation in complex human-robot collaborative tasks. In **Proc. RO-MAN**, pp. 1119–1126, 2020.
 - [4] Francisco Cruz et al. Explainable robotic systems: Understanding goal-driven actions in a reinforcement learning scenario. **Neur. Compu. and App.**, pp. 1–18, 2021.
 - [5] S. Marchesi et al. Don’t overthink: fast decision making combined with behavior variability perceived as more human-like. In **Proc. RO-MAN**, pp. 54–59, 2020.
 - [6] Maartje MA de Graaf, Bertram F Malle, Anca Dragan, and Tom Ziemke. Explainable robotic systems. In **Proc. HRI**, pp. 387–388, 2018.
 - [7] Wenxuan Mou, Martina Ruocco, Debora Zanatto, and Angelo Cangelosi. When would you trust a robot? a study on trust and theory of mind in human-robot interactions. In **Proc. RO-MAN**, pp. 956–962, 2020.
 - [8] Connor Esterwood and Lionel P. Robert. Do you still trust me? human-robot trust repair strategies. In **Proc. RO-MAN**, pp. 183–188, 2021.
 - [9] Ilenia Cucciniello et al. Validation of robot interactive behaviors through users emotional perception and their effects on trust. In **Proc. RO-MAN**, pp. 197–202, 2021.
 - [10] Glenda Hannibal, Astrid Weiss, and Vicky Charisi. “the robot may not notice my discomfort” – examining the experience of vulnerability for trust in human-robot interaction. In **Proc. RO-MAN**, pp. 704–711, 2021.
 - [11] Arianna Pipitone and Antonio Chella. What robots want? hearing the inner voice of a robot. **iScience**, Vol. 24, No. 4, 2021. DOI: 10.1016/j.isci.2021.102371.
 - [12] David Gunning, Mark Stefik, Jaesik Choi, Timothy Miller, Simone Stumpf, and Guang-Zhong Yang. Xai—explainable artificial intelligence. **Science Robotics**, Vol. 4, No. 37, p. eaay7120, 2019.
 - [13] Rossitza Setchi, Maryam Banitalebi Dehkordi, and Juwairiya Siraj Khan. Explainable robotics in human-robot interactions. **Procedia Computer Science**, Vol. 176, pp. 3057–3066, 2020.
 - [14] Qingxiaoyang Zhu, Vittorio Perera, Mirko Wächter, Tamim Asfour, and Manuela Veloso. Autonomous narration of humanoid robot kitchen task experience. In **Proc. Humanoids**, pp. 15–17, 2017.
 - [15] Zhao Han, Elizabeth Phillips, and Holly A. Yanco. The need for verbal robot explanations and how people would like a robot to explain itself. **ACM Trans. HRI**, Vol. 10, No. 4, pp. 1–42, 2021.
 - [16] Marcella Cornia, Lorenzo Baraldi, and Rita Cucchiara. Smart: training shallow memory-aware transformers for robotic explainability. In **Proc. ICRA**, pp. 1128–1134. IEEE, 2020.
 - [17] Koichiro Yoshino et al. Caption generation of robot behaviors based on unsupervised learning of action segments. In **Convers. Dialog. Sys. for the Next Decade**, pp. 227–241. Springer, 2021.
 - [18] Pradip Pramanick, Chayan Sarkar, and Indrajit Bhat-tacharya. Your instruction may be crisp, but not clear to me! In **Proc. RO-MAN**, pp. 1–8, 2019.
 - [19] Fatai Sado, C Kiong Loo, Matthias Kerzel, and Stefan Wermter. Explainable goal-driven agents and robots—a comprehensive review and new framework. **arXiv preprint arXiv:2004.09705**, Vol. 180, , 2020.
 - [20] Maria Fox, Derek Long, and Daniele Magazzeni. Explainable planning. **arXiv preprint arXiv:1709.10256**, 2017.
 - [21] Tatsuya Sakai, Kazuki Miyazawa, Takato Horii, and Takayuki Nagai. A framework of explanation generation toward reliable autonomous robots. **Adv. Robot.**, Vol. 35, No. 17, pp. 1054–1067, 2021.
 - [22] 湯口彰重, 河野誠也, 石井カルロス寿憲, 吉野幸一郎, 川西康友, 中村泰, 港隆史, 斉藤康己, 美濃導彦. ぶつくさ君: 自身の外界認識と内部状態を言語化するロボット. **日本ロボット学会誌**, Vol. 40, No. 10, pp. 932–935, 2022.
 - [23] Aaqib Tabrez and Bradley Hayes. Improving human-robot interaction through explainable reinforcement learning. In **Proc. HRI**, pp. 751–753. IEEE, 2019.
 - [24] Hooman Hedayati, Mark D. Gross, and Daniel Szafir. What information should a robot convey? In **Proc. IROS**, pp. 6209–6216, 2021.
 - [25] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. Interactive robots as social partners and peer tutors for children: A field trial. **HRI**, Vol. 19, No. 1-2, pp. 61–84, 2004.
 - [26] Iolanda Leite, Carlos Martinho, Andre Pereira, and Ana Paiva. As time goes by: Long-term evaluation of social presence in robotic companions. In **Proc. RO-MAN**, pp. 669–674. IEEE, 2009.
 - [27] Kazunori Komatani, Shinichi Ueno, Tatsuya Kawahara, and Hiroshi G Okuno. User modeling in spoken dialogue systems for flexible guidance generation. In **Proc. EUROSPEECH**, pp. 745–748, 2003.
 - [28] Taku Kudo and John Richardson. Sentencepiece: A simple and language independent subword tokenizer and detokenizer for neural text processing. **arXiv preprint arXiv:1808.06226**, 2018.
 - [29] Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier, and Michael Auli. fairseq: A fast, extensible toolkit for sequence modeling. In **Proc. NAACL-HLT**, 2019.
 - [30] 趙天雨, 沢田慶. 日本語自然言語処理における事前学習モデルの公開. **人工知能学会研究会資料 (SLUD)**, Vol. 93, pp. 169–170, 2021.
 - [31] Thomas Wolf et al. Transformers: State-of-the-art natural language processing. In **Proc. EMNLP**, pp. 38–45, Online, October 2020. Association for Computational Linguistics.
 - [32] Tatsuya Nomura et al. Measurement of negative attitudes toward robots. **Interaction Studies**, Vol. 7, No. 3, pp. 437–454, 2006.
 - [33] Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. Assessing acceptance of assistive social agent technology by older adults: the almere model. **Interna. jour. of soc. robot.**, Vol. 2, No. 4, pp. 361–375, 2010.
 - [34] 川西康友, 吉野幸一郎, 岡留有哉, 中村泰, 石井カルロス寿憲, 港隆史, 斉藤康己, 美濃導彦. ぶつくさ君 2号: 自律移動ロボットによる生活環境の認識. 第25回画像の認識・理解シンポジウム (MIRU), 2022.

付録

A 実験で使⽤したロボット

使⽤する⾃律移動型対話ロボットプラットフォーム「ぶつくさ君 [22, 34]」である。「ぶつくさ君」の外観を図 2 に⽰す。

ぶつくさ君は内部状態として観測・記憶、案内タスクを実⾏するという意図を持ち、観測や記憶を意図で処理して移動や発話を行う。システム全体は Robot Operating System (ROS) で管理されており、各外界観測はそれぞれ個別の ROS ノードとして処理され、記憶は MongoDB を⽤いて管理されている。

ぶつくさ君は、物体認識、⼈物認識、環境認識、音声認識、場所認識の外界観測を行い、その認識結果をいつ、どこで⾏ったのかという、場所・時間の情報と紐づけて記憶することができる。また、動作機能として移動、モーション、音声の生成ができ、あらかじめ設定された部屋の地図上の特定の場所に移動や、任意のセリフを発話することができる。ぶつくさ君が動作する部屋の間取り図を図 3 に⽰す。

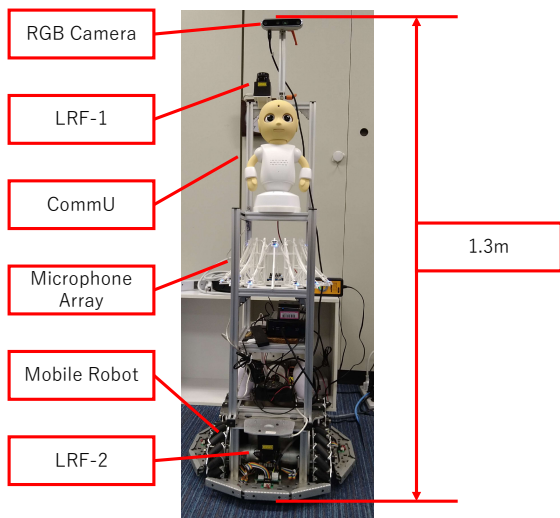


図 2 ロボットプラットフォームぶつくさ君

B 言語化規則

ロボットの発話は表 2 の規則によって生成される。デフォルトとして、内部状態は説明の充分性を担保するために多少冗長と⾐じられてもすべて話したほうが良いという先行研究 [22] の知見に基づき、挨拶、⾏動、⾏動の理由、音声・場所認識の報告、音声・場所認識の記憶の報告を発話するように設定した。シナリオのターンテークはユーザの発話

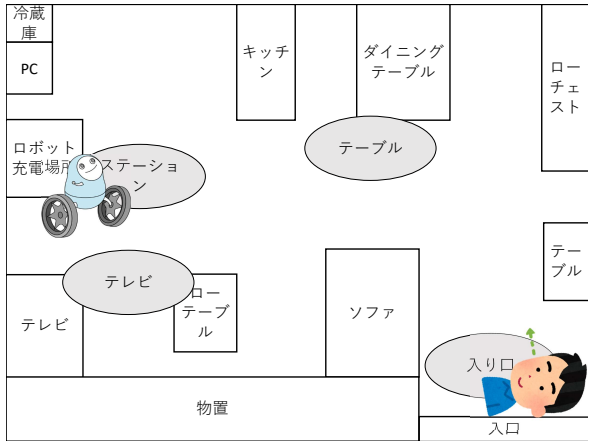


図 3 実験室の間取り図

に対し特定のトリガーワードをロボットが音声認識することによって進⾏する。

表 2 案内タスクでの言語化規則。[場所]、[物体]、[音声]、[時間] は認識結果。

spk	from	dest	dialogue
rob	始点	始点	入り口の近くに移動するよ。
rob	始点	入口	入り口の近くに移動するのは玄関に人がいると判断したからだよ。
rob	入口	入口	入り口の近くに着いたよ。
rob	入口	入口	入り口の近くに着いたことを記憶したよ。
rob	入口	入口	こんにちは。
usr	入口	入口	(こんにちは)
rob	入口	入口	[音声] という音声を認識したよ。
rob	入口	入口	[音声] という音声を認識したことを記憶したよ。
rob	入口	入口	何か探しものはありますか？
rob	入口	入口	何か探しものはありますか？と聞くのは、入り口の近くにいる人は探しものがあると判断したからだよ。
usr	入口	入口	([物体]) を探しています。
rob	入口	入口	[音声] という音声を認識したよ。
rob	入口	入口	[音声] という音声を認識したことを記憶したよ。
rob	入口	入口	[音声] という音声からあなたは [物体] を探していると思うよ。
rob	入口	入口	そのように判断したのは、[音声] という音声の中に、「[物体]」という単語を認識したからだよ。
rob	入口	入口	[物体] は [場所] の近くにあると思うよ。
rob	入口	入口	そのように判断したのは、時間、時間に [場所] の近くで物体、物体を認識したことを記憶しているからだよ。
rob	入口	入口	[場所] の近くまで案内しようか？
rob	入口	入口	[場所] の近くまで案内しようか？と聞くのは、案内が必要かどうか確認するためだよ。
usr	入口	入口	(お願い)
rob	入口	入口	[音声] という音声を認識したよ。
rob	入口	入口	[音声] という音声を認識したことを記憶したよ。
rob	入口	入口	[場所] の近くに移動するよ。
rob	入口	[場所]	[場所] の近くに移動するのは、[音声] という音声から案内が必要だと判断したからだよ。
rob	[場所]	[場所]	[場所] の近くに着いたよ。
rob	[場所]	[場所]	[場所] の近くに着いたことを記憶したよ。
usr	[場所]	[場所]	(ありがとう)
rob	[場所]	[場所]	[音声] という音声を認識したよ。
rob	[場所]	[場所]	[音声] という音声を認識したことを記憶したよ。
rob	[場所]	[場所]	ここから離れるよ。
rob	[場所]	[場所]	ここから離れるのは、あなたがもう案内を必要としないと判断したからだよ。
rob	[場所]	始点	さようなら
usr	始点	始点	(さようなら)