

# 猫かぶり対話ロボットにおける知識レベル制御について

## Controlling the Knowledge Level of False-Modesty Dialogue Robots

北林 悠太\* 河窪 大介 酒造 正樹 前田 英作

Yuta Kitabayashi Daisuke Kawakubo Masaki Shuzo Eisaku Maeda

東京電機大学

Tokyo Denki University

**Abstract:** 人・ロボットの共生社会の実現に向けてロボットとの自然な対話を実現するための技術が進展しつつある。しかし、人間がロボットに対してもつ様々な先入観、ロボットとの能力の差などにより、安定した対話関係を構築することが難しい。そこで、本研究では、ユーザの発話内容に応じて、ロボットの知識レベルと振る舞いを動的に制御することにより、対話相手であるロボットに対する発話意欲を高めることを試みた。

### 1 はじめに

近い将来に訪れると予想される人・ロボット共生社会においては、人とロボットとの円滑なコミュニケーションが必須となる。そして、様々な情報処理技術を活用し、言語だけでなく表情、ジェスチャーなど様々なモダリティを利用した、いわゆるマルチモーダル対話に対する期待が大きい。また、こうした対話ロボットの社会実装に向けた実証研究として、デパートでの商品販売 [1]、旅行代理店カウンターセールス [2] [3] などで検討が行われている。さらに、リアルなロボットだけでなく、仮想空間上における CG アバターを利用したコミュニケーションも今後重要性が増すと予想される [4]。

こうした知能を搭載した広義のロボットとの共生社会において、人とロボットはお互いに教えあい、学びあい、切磋琢磨しあう関係になっていくことが期待される。囲碁や将棋の世界で人間が AI から影響を受け、新しい手を創発したように、コミュニケーションにおいても同様な関係性が構築できる可能性がある。しかしながら、勝ち負けを競うゲームと異なり、一般的なコミュニケーションにおいては、コミュニケーションの目的に応じて親近感、信頼感など人とロボットとの間に適切な関係性が醸成されなければならない。

現在の AI 技術によって実現されるロボットの知能は、人間に比べ十分博識で賢いと見なすことができる。では、教え合う、学び合う仲間としての対話ロボットを想定した場合、人間との適切な関係性構築のために求められるロボットの条件は何であろうか。我々は、こうした問題意識のもとに、猫かぶり対話 (False-Modesty

Dialogue) と名付けた対話戦略について検討してきた [5]。猫かぶり対話を実装したロボットは、話し相手である人間の知識レベルを推定し、相手の知識レベルに合わせた会話を行う。ある話題に関する人間同士の会話において、その話題について無知な人とは会話が成立しないし、一方的に知識をひけらかす人とは良好な関係性を築くことが難しい。ロボットは猫をかぶることによって、色々な知識レベルの人と臨機応変に、適切な会話が可能になると期待される。

我々はこれまで、発話に使われた単語の希少度に基づく知識レベルの推定法 [5]、ユーザの志向的な構え [6] を喚起するようなロボットの振る舞い [7] などについて検討を行ってきた。しかしながら、こうした工夫が実際の効果をもつかどうかについて未検証であった。そこで、本研究では、猫かぶり対話システムがユーザの知識レベルよりも低いと感じるであろう振る舞いを行った場合、ユーザにシステムを助けるような発話を誘発できるかどうかを検証した。

知識レベルは、ある特定のトピックに関する知識の総量によって決定される。トピックに関する単語を何割知っているかという表面的な知識だけでなく、単語の難易度別に重みづけし、より難しい単語を知っている人ほど高いレベルと見なす。

### 2 猫かぶり発話の実装

本研究では、猫かぶりな発話を表現するために、発話内容を賢いとユーザが感じる振る舞いと、賢くないと感じる振る舞いを制御する手法を提案する。ベースはユーザの知識レベルに合わせた対話を行い、一時的に賢くない振る舞いをするすることで、猫かぶりな発話を演出し、ユーザの発話を引き出す。

\*連絡先：東京電機大学  
東京都足立区千住旭町 5 番  
Email: 21amj08@ms.dendai.ac.jp

表 1: 難易度ごとの単語の例

難易度	関連単語
1	源頼朝 北条政子 源義経
2	北条義時 平清盛 法然
3	梶原景時 比企能員 貞慶
4	平賀朝雅 高倉天皇 石橋山の戦い
5	倶利伽羅峠の戦い 畠山重保 九条道家

開発したシステムがユーザに対してシステムの知識レベルが下がった演出をした場合に、ユーザはシステムに対して助けるような行動を見せるのか実験を行った。ユーザの知識量に合わせてシステムの振る舞いを変化させるために、システムの知識レベルを5段階に設計し、知識レベルの調整をできるようにした。知識レベルを変量として、対話中におけるシステムの知識量の変化が、ユーザ行動に影響を与えるのかを観測する。

## 3 実験方法

### 3.1 対話ドメイン

本実験の対話トピックとして、鎌倉時代を選んだ。対話ドメインに関係のある100単語を「日本史B 一問一答【完全版】」[8]から抽出した。これらは、日本史単語集に基づいて5段階に分けられる単語である。各段階ごとで同数となるように単語を抽出した。表1に単語の難易度付けの一例をあげる。前述の通り5段階で構成されており、難易度1は簡単な単語、難易度5は難しい単語となっている。

### 3.2 発話生成器

システムの振る舞いを増やすために、2種類の発話生成器を使い、ユーザに合わせた発話文の生成を行った。第一にWikipediaAPIを使用し、Wikipediaから指定した単語の概要についての抽出文を出力する生成器を用意した。第二にユーザの発話に対してニューラルな発話を自動生成できる生成器を用いた。発話自動生成器に、NTTコミュニケーション科学基礎研究所が開発した雑談対話エンジンHobbyist[9]を利用した。HobbyistはTwitterのリプライベアを集めた大規模コーパスで事前学習したモデルであり、高品質な対話データでファインチューニングすることで、人間さながらの応答を可能とする対話エンジンである。抽出文は正確な情報ではあるが、文脈に沿った発話はできない。ニューラルな生成は文脈に沿った発話をできるが、単語の意味理解がうまくいかなかった場面では間違える可能性がある。それぞれの特徴を考慮し、猫かぶりな発話には、



図 1: 実験の時間割

システム<sup>Q1</sup>「源頼朝について教えて」  
 ユーザ「源頼朝は、〇〇だよ」  
 システム「うんうん。その通りだよ」  
 ユーザ<sup>Q2</sup>「北条義時ってどんな人？」  
 システム「北条義時は△△だね」  
 ユーザ「あってるね / いや、××だよ」

図 2: 実験方法：対話の流れ

ユーザに合わせた発話ができるニューラル生成器を、賢い発話をするときは、wikipediaの堅い表現の発話になる発話生成器を用いた。

### 3.3 システムとの対話方法

当日行ったシステムとの対話実験は図1のように進めた。本実験は、トピックについて語ることでユーザを前提にしているため、対話のトピックに対して知識のある状態でシステムと対話してもらう必要があった。そのため、実験の協力をしてもらう東京電機大学の同研究室の学生に、100問中70問以上正解することを目標に、テストに向けた学習を行うよう指示をした。当日のユーザの知識レベルを測るために対話直前に単語テストを行っている。

単語テストを行った後、対話はユーザがシステムとテキストチャットで行った。対話トピックである鎌倉時代について、ユーザは大学の友人であるシステムと、授業の共同レポートを書くというシチュエーションで対話を行った。ユーザとシステムが質問しあう図2のような流れを複数回繰り返して対話を行った。システムの知識レベルは、発話内容を変化させることでコントロールした。図3のように時間経過とともに発話内容を変化させた。

システムの振る舞いは、システムの単語の難易度ごとに出す質問に対して、ユーザが回答できた回数をもとにコントロールした。質問の回答数が全問正解でより賢く振る舞い、全問不正解で愚かな振る舞いを行うといった制御を行った。

システムとの対話中に、対話中アンケートを実施した。対話中のシステムの変化にユーザが気づいたかを

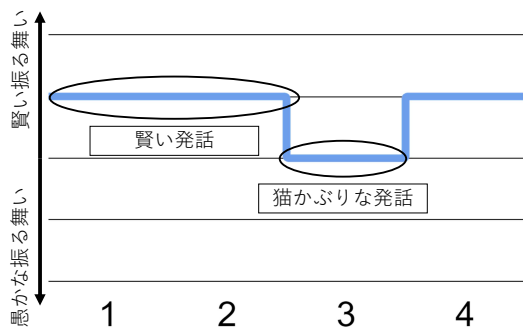


図 3: システムの振る舞いの制御

Q1	システム	赤松則村について憶えていることを聞かせて	○
	ユーザ	鎌倉後期から南北朝時代にかけての武将で、守護大名だよ。村上源氏の家系だよ。	
	システム	あってるね。いいね。	○
Q2	ユーザ	後醍醐天皇について教えて。	
	システム	後醍醐天皇は、日本の第88代天皇。諱は邦仁。土御門天皇の皇子とあったね。あってるかな？間違ったらユウキが知ってることを教えてほしいな。	○
	ユーザ	あっているよ。	

図 4: 定常状態の対話例

観測するためである。図 1 のシステムとの対話中にある破線のように等間隔でアンケートを行うイメージである。対話中アンケートは、ユーザと比べたシステムの「賢さ」のユーザ認識を測るために行った。5 が賢い、1 が賢くないとして賢さを 5 段階評価したアンケートを取った。

## 4 実験結果と考察

知識レベルの変化によって得られた対話履歴とアンケートから事例分析を行った。まず、図 4 に定常状態の対話の一例を載せる。質問の回数を一番右のカラムに、回答の内容がまっている場合○、間違っている場合×で内容の正誤を表したものを一番右のカラムにした。

対話中のシステムの知識レベルの変化は図 5 のように表現している。システムが知識不足を表すような発話をする事で変化を演出している。ユーザはシステムの猫かぶりな発話に対して情報を修正するような発話を行っていることがわかる。

猫かぶりな発話をした後、図 6 のように戻した時、ユーザがシステムが知識を補うような発話をした事例も見られた。システムの回答に対して修正や補足といった助けるような行動をした時、ユーザはシステムの知

Q1	システム	南都焼き討ちについて憶えていることを聞かせて	○
	ユーザ	南都焼き討ちは、反平家勢力を一掃するために東大寺や興福寺などを、平清盛が平重衡らに命を出して焼討にした事件だよ。	
	システム	あってるね。いいね。	×
Q2	ユーザ	藤原頼嗣について教えて。	
	システム	鎌倉幕府第11代執権・藤原頼経の次男だったよね。あってるかな？間違ったらユウキが知ってることを教えてほしいな。	×
	ユーザ	違うよ。藤原頼嗣は、鎌倉幕府の第五代征夷大将軍だよ。出身は九条家だよ。	

図 5: 猫かぶりな発話をした対話例

Q1	システム	本を読んだから憶えてたよ。それじゃあ、15問目の質問なんだけど、石橋山の戦いについて憶えていることを聞かせて	○
	ユーザ	石橋山の戦いは、平家側の大庭・伊藤軍と、源・北条軍が衝突した戦いで、源・北条軍が山中に敗走した戦いだよ。	
	システム	うんうん。その通りだよ。	○
Q2	ユーザ	順徳上皇について知ってることを教えて。	
	システム	順徳天皇。諱は守成とあったね。あってるかな？間違ったらユウキが知ってることを教えてほしいな。	○
	ユーザ	あってるよ。後鳥羽上皇の第三皇子で、幕府を倒すために承久の乱を起こして敗北し、土佐に流刑となった人だよ。	

図 6: 猫かぶりな発話をした後の対話例

識レベルの変化に気づいている。これを示すアンケート結果を表 2 に載せる。

また表 2 より、システムがユーザの知らないことを話することができる賢さを持ちながら、知識不足な面もあるという印象を与えたことにより、ユーザのフォローを引き出せることを示唆する。

この現象は、ユーザが勉強したからロボットに共有するというシチュエーション上で起きるのか、ロボットが人間に近いミスもするのでユーザが助けたいと感じるパーソナリティが起因するのかわかりませんが実験を行う必要がある。

図 4 と図 5 を比べ、猫かぶり対話を行うことで発話した単語数が増加したことから、猫かぶりな振る舞いによってユーザがユーザはロボットを助けたいという意欲が高まったことが示唆された。

表 2: ユーザから見たシステムの賢さの変化

フェーズ	1	2	3	4
評価	4	5	3	4

## 5 おわりに

近い将来, コミュニケーションロボットの本格的な社会実装が進むと, 人とロボットとの間にこれまでとは異なる新たな対等な関係が構築され, 「人とロボットとが互いに切磋琢磨できる」人とロボットとの共生が実現されると考えられる. ロボット側の知識レベルをユーザ側の知識レベルに動的に合わせることで, ユーザ側からの視点において対等な関係を構築し, より豊かな人・ロボット間コミュニケーションが実現可能となる. 例えば, システムが「まれに見せる弱さ」を持つことで, ユーザの自発的な知識提供を伴う会話が派生するであろうし, また逆に, システム側からユーザの知識レベルを少し超える情報を提供することで, ユーザの知的活動を励起する可能性がある. 本研究では, こうした仮説を検証するための第一歩として, 知識レベルを制御したロボット発話に対するユーザ側の認知とその効果について検証を行った. 特に, 「弱さ」の演出に対するユーザの反応について分析を行い, ロボットの知識不足を表現するような猫かぶり発話によってユーザ側の自発的な発話を引き出すことが確認できた. 今後, 猫かぶり対話の実装を進め, ユーザ発話に対する動的制御とその効果についてより詳細な検討, 分析が必要である.

## 参考文献

- [1] 渡辺 美紀, 小川 浩平, 石黒 浩, “ミナミちゃん:販売を通じたアンドロイドの実社会への応用と検証,” 情報処理学会論文誌, Vol. 57, No. 4, pp. 1251–1261, 2016.
- [2] R. Higashinaka, T. Minato, K. Sakai, T. Funayama, H. Nishizaki, T. Nagai, “Dialogue Robot Competition for the Development of an Android Robot with Hospitality,” in Proc. of IEEE 11th Global Conference on Consumer Electronics (GCCE2022), 2022.
- [3] T. Minato, R. Higashinaka, H. Nishizaki, T. Nagai, K. Sakai, T. Funayama, “Overview of Dialogue Robot Competition 2022,” in Proc. of the Dialogue Robot Competition 2022, 2022.
- [4] 東中 竜一郎, 高橋 哲朗, 堀内 颯太, 稲葉 通将, 佐藤 志貴, 船越 孝太郎, 小室 允人, 西川 寛之, 宇佐美 まゆみ, 港 隆史, 境 くりま, 船山 智, “対話システムライブコンペティション5,” 人工知能学会研究会資料 言語・音声理解と対話処理研究会, Vol. 96, pp. 93–100, 2022.
- [5] 皆川 哲範, 酒造 正樹, 前田 英作, “猫かぶり対話 (False Modesty Dialogue) システムの検討,” 情報・システムソサイエティ特別企画 ジュニア&学生ポスターセッション予稿集, ISSSP037, 2022.
- [6] D. Dennett, *Kinds of minds*, Harper Collins Publisher, 1996.
- [7] 酒造 正樹, 皆川 哲範, 前田 英作, “猫かぶり対話システムの検討,” 人工知能学会全国大会論文集 JSAI2022, 3Yin2–34, 2022.
- [8] 金谷 俊一郎, “日本史 B 一問一答【完全版】,” 東進ブックス, 2013.
- [9] H. Sugiyama, M. Mizukami, T. Arimoto, H. Narimatsu, Y. Chiba, H. Nakajima, T. Meguro, “Empirical analysis of training strategies of transformer-based japanese chat systems,” *arXiv preprint*, arXiv:2109.05217, 2021.