

「空気を読む」車載 AI：言ってほしくないことは言わず、 言ってほしいことを言う対話情報管理システム

Socially Intelligent Automotive AI: A Dialogue Information Management System That Selects What Should and Should Not Be Said Based on Situational Context

前田 菜月^{1*} 飯田 愛結¹ 奥岡 耕平¹ 坂本 忠司² 雪田 幸宏² 大澤 正彦¹

Natsuki Maeta¹, Ayu Iida¹, Kohei Okuoka¹,

Tadashi Sakamoto², Yukihiro Yukita², Masahiko Osawa¹

¹ 日本大学

¹ Nihon University

² トヨタ自動車株式会社

² Toyota Motor Corporation

Abstract: 本研究では、車載 AI の所有者である運転者が同乗者に知られたくない情報を安全に管理しつつ、同乗者と理想の対話が行われるよう支援する「空気を読む」対話情報管理システムを提案する。提案システムでは、車載 AI と運転者の二者間と、同乗者を含む複数者間の対話ログを分離する。さらに車内での対話であるという特徴を活かし、運転者と車載 AI が同乗者に開示可能な情報をあらかじめ精査し、自然に表出する方法を検討する。

1 はじめに

本研究の目的は、単なる便利な道具としてではなく運転手のパートナーとして信頼される存在となる車載 AI の実現である。自動運転技術や大規模言語モデル (LLM) などの自然言語処理技術の発展に伴い、車両と人とのインタラクションにおいて車載 AI が果たす役割が重要性を増している。本研究では、車両に搭載された、音声や言語を介して乗員と対話しながら道案内などの情報提供や運転操作の支援などを行うシステムを車載 AI と呼称する。車載 AI は、運転状況や車内環境といった車ならではの文脈を踏まえた対話戦略が求められる。

ここで、車内という空間は密室のプライベートな空間であると同時に、家族や友人、時には仕事相手などの他者を招き入れてコミュニケーションを行う社会的空間にもなる。そのため車載 AI は、1 対 1 のコミュニケーションだけではなく、複数人でのコミュニケーションを扱えることが求められる。特に、プライベートな対話とパブリックな対話が同じ空間で発生し得る車載 AI では、話者構成に応じた情報の扱いに注意が必要である。例えば、運転手と車載 AI の二者間で行われた対話内容

が他の同乗者が居る状況下で車載 AI から発話されることで、運転手にとって不利益な発話に繋がる可能性がある。

既存研究では、LLM を用いた対話システムにおいてユーザから得られた機密情報を適切に管理するさまざまな手法が提案されている。例えば Hemken らの研究では、機密情報のアクセス要求に関するデータセットを用いて LLM をファインチューニングすることで、最大 98% という非常に高い精度で判断できるようになることを報告している [1]。しかし、機密情報の漏洩は 1 度でも致命的な問題となるため、実応用の場面では話者構成毎に対話データを別に管理するなど、構造的な情報分断を基本とする設計原理が採用されることが多い。

一方でユーザ体験を高める観点では、ユーザの情報を蓄積してパーソナライズすることや柔軟な応答生成を可能にすることが車載 AI の有用性に繋がる。そのため、話者構成ごとに情報を完全に切り離してしまうと、これらの恩恵が失われ、システムの有用性とのトレードオフが生じる。そこで本研究では、ユーザが他者に言ってほしくない情報は言わずに、ユーザが使用を認める情報は活用して支援する空気を読んだ対話を行う「空気を読む」車載 AI の実現を目指す。

本研究では「空気を読む」車載 AI の基盤技術として、同乗者の構成に応じて対話ログを管理する情報管理機

*連絡先：日本大学文学部
〒156-8550 東京都世田谷区桜上水 3-25-40
E-mail: chna23084@g.nihon-u.ac.jp

能と、運転手に対話で得られた情報の開示可否を確認する開示確認機能を有する対話情報管理システムを提案する。LLMと認知モデルを統合した対話アーキテクチャを拡張して提案する対話情報管理システムを実装し、2つのシナリオを用いて実装したシステムが設計通りに機能するか簡易的な動作検証を行った。加えて、実際の車両への搭載を見据え、車載環境下におけるシステムの検証を容易に行うためのUnityを用いた3Dシミュレーション環境を構築した。シミュレーション環境を用いることで、車に搭載された様々なセンシング機能や周辺の状態の活用が求められる実環境下を想定した検証が簡易に実現できる。

2 背景：「空気を読む」車載 AI

本章では、本研究の背景として、車内空間における「空気を読む」車載 AI の必要性和実現可能性について説明する。

2.1 車内空間における「空気を読む」車載 AI の必要性

車両は閉鎖空間であり、外部環境と分断することで、プライベートな空間を構築することができている。そのため、一人で運転している最中に周囲の人間に盗み聞きされる心配がなく、運転者と車載 AI は完全にプライベートな空間で会話することができる。したがって将来的に車載 AI が運転者の信頼を得て、重要なパートナーとなった場合、多岐にわたる情報を車載 AI に話す可能性が高い。さらに昨今の対話型 AI に対する人々の反応を鑑みれば、運転者は車載 AI に対して、会話した履歴を維持・学習し、運転手のパーソナルな存在になることを求めると考えられる。

しかしここには、同乗者への情報漏洩という重大なリスクが潜んでいる。車内には、運転手以外の同乗者が存在する可能性がある。実際に情報漏洩すること自体の問題があるだけでなく、情報漏洩するかもしれないという不安感は、運転手から車載 AI への自己開示を妨げ、信頼構築を実現しにくくなると予想される。したがって、車載 AI に特化した対話情報管理システムが必要となる。

ここで求められる対話情報管理システムは、単に運転手との会話を一切漏洩しない振る舞いをするのでは不十分である。運転手が車載 AI をパーソナルエージェントとして信頼するためには、様々な対話の履歴を踏まえた応答を実現できる必要がある。つまり、言ってほしくないことは言わず、言ってほしいことを言う、「空気を読む」車載 AI が必要である。

2.2 車内空間における「空気を読む」車載 AI の実現可能性

車内空間において「空気を読む」車載 AI が必要であると述べた一方で、車内空間は AI の実現可能性が高い環境ともいえる。その理由として主に2点が挙げられる。

1点目は、会話グループの特定が容易であるためである。ここで会話グループとは、ある会話への参加者の集合を意味する。会話内容の情報漏洩の問題は、ある会話グループ内で得た情報を、別の会話グループで発話してしまうことによって生じる。

車内空間において、会話グループへの参加者は、多くの場合において同乗者全員という仮定において問題がない。なぜなら、一般的な乗用車において、同乗者に対して知られたいくない情報が発話されることは考えにくいためである。一方車以外の環境においては会話グループへの参加者の同定は一定の誤認識のリスクがある。例えばオープンな環境では、検出した人間全員を会話グループへの参加者と認識することは妥当ではない。なぜなら、同一空間において、複数の会話グループが形成されることや、会話グループに参加していない人がいるケースがあるためである。

また、車両は車種によって乗車定員が決まっており、会話グループの上限も決まっている。さらに商用車でなければ、同乗者になりうる人も限定的であり、個別に認識してデータ管理をすることも現実的と考えられる。更に、車両内にカメラやマイク、シートセンサーなど様々なセンサーを搭載することができる上に、同乗者の位置もほとんど規定された位置であると考えられるため、同乗者の認識が成功しやすい。

2つ目の車内空間において空気を読む会話の実現しやすい理由は、運転者と車載 AI が確実に二人だけの対話を十分担保できるためである。そもそも、運転手が車載 AI に対して「言ってほしいこと」「言ってほしくないこと」を認識することは難しい問題である。しかしながら、車載 AI であれば運転手と車載 AI の2者間対話は日常の中に一定の時間が担保されていると考えられる。したがって、2者間対話を通じた綿密な情報開示可否のすり合わせなどを行うことができると考えられる。

以上を踏まえて、次章では、車載 AI の特性を活かし、「会話グループの識別が確実にできること」「2者間対話によって十分な情報共有が運転者と車載 AI の間で可能であること」を前提とした、車載 AI の対話情報管理システムを提案する。

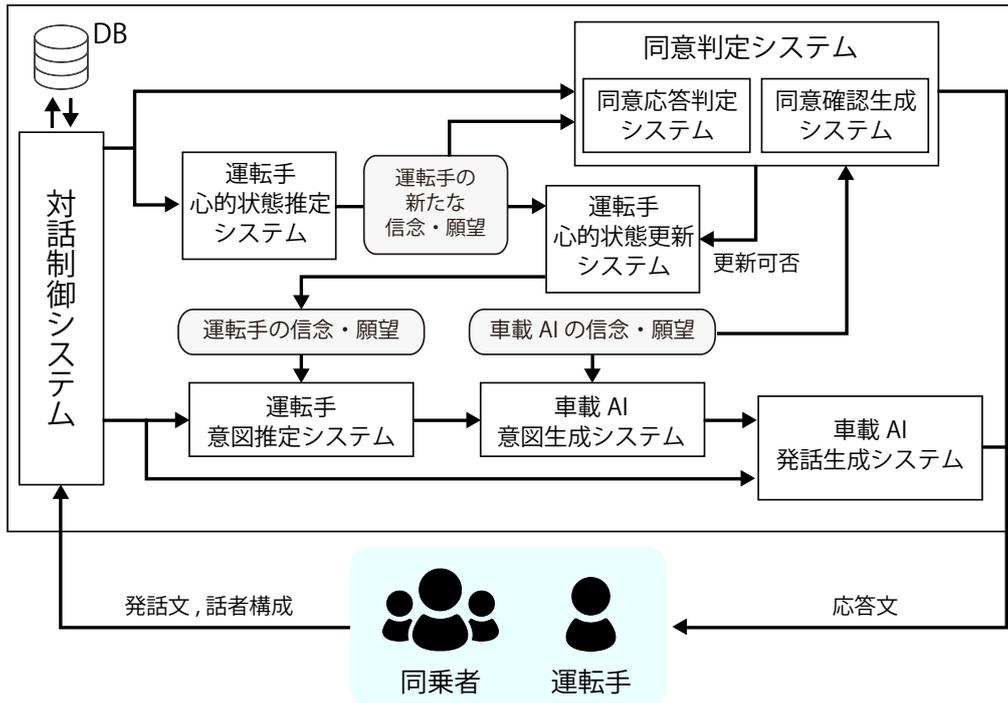


図 1: 対話情報管理システムの構成図

3 対話情報管理システム

本研究では「空気を読む」車載 AI の実現に向けた基盤技術として、対話情報管理システムを提案する。提案システムは、搭乗者構成に応じて対話ログを分割して保存する情報管理機能と、必要に応じて運転手との二者間対話で得られた情報を他者間の対話に開示する可否を確認する開示確認機能を有する。情報管理機能によってユーザが望まない情報の漏洩を構造的に防ぐだけでなく、開示確認機能によって必要に応じて二者間対話で得られた情報を他者間の対話に活用することで、言ってほしくないことは言わず、言ってほしいことを言う「空気を読む発話」が実現できる。提案システムの構成を図 1 に示す。

提案システムは、先行研究で [2] 提案する、Belief-desire-intention(BDI) モデルベースの認知アーキテクチャと大規模言語モデルを統合する手法を拡張して実装した。以降の節で、各システムの詳細について述べる。なお、本研究では大規模言語モデルには OpenAI の GPT-4o(バージョン:gpt-4o-2024-05-13) を用いた。

3.1 対話制御システム

対話制御システムでは、提案システムにおけるログデータの管理を行う。ログデータの管理では、入力された発話文を話者ごと、また話者構成ごとに分けてデー

タベースに保存する。その後、入力された話者ごと、話者構成に沿った対話履歴をデータベースから読み込み、必要に応じて各システムに出力する。

同乗者が乗った場合の 3 者間対話では、先行研究で [2] 提案された対話認知モデルに従って発話を生成させる。この時、運転手からの許可を得て信念と願望が更新されていた場合、それらの情報を踏まえて発話が生成される。反対に、更新されていなかった場合は事前に設定されていた信念と願望のみを用いて発話生成のプロセスが行われる。

3.2 運転手心的状態推定システム

運転手心的状態推定システムは、運転手との過去の対話履歴を入力に、運転手の信念および願望を大規模言語モデルを用いて推定する。以下にそのプロンプトについて説明する。

はじめにインタラクションの系 (A.1) とシステムが持つ内部表現 (A.2) を説明するプロンプトを与える。ここで心的状態推定システムが持つ内部表現は、「他者の信念」「他者の願望」「他者の意図」である。続いて信念・願望・意図に関するプロンプト (A.3) を与えたのち、以下のシステム説明を与える。

心的状態推定システムについて説明します。心的状態推定システムは、「他者の発話から、他者の信念と願望を推論する」システムです。入力として与えられた、「他者の発話」から、矛盾や違和感のない「他者の信念」「他者の願望」を「推論」してください。

最後に入出力を指定するプロンプト (A.4) を与えて、出力を得る。ここで心的状態推定システムの入力は「他者の発話」であり、出力は「他者の信念」と「他者の願望」である。

3.3 同意判定システム

同意判定システムには、同意確認生成と同意応答判定の2つのシステムから構成される。これらは、2者間対話で得られた情報(詳細は3.2項を参照)を他者間の対話に開示する可否を確認する開示確認機能として動作する。以下に、各システムについて順に説明する。

3.3.1 同意確認生成システム

同意確認生成システムは、運転手心的状態推定システム(3.2)から出力される運転手の信念および願望と、車載AI(ナビゲーションシステム)の信念および願望を入力し、開示の可否を確認する応答文を大規模言語モデルを用いて生成する。以下にそのプロンプトについて説明する。

はじめにインタラクションの系(A.1)とシステムが持つ内部表現(A.2)を説明するプロンプトを与える。ここで同意確認生成システムが持つ内部表現は、「自己の信念」「他者の信念」「自己の願望」「他者の願望」「自己の意図」「他者の意図」である。続いて信念・願望・意図に関するプロンプト(A.3)を与えたのち、以下のシステム説明を与える。

同意確認生成システムについて説明します。同意確認生成システムは、「仮説として受け取った他者の信念と願望を、情報として保持していいか確認するための質問を生成する」システムです。入力として与えられた、「他者の信念」「他者の願望」「自己の信念」「自己の願望」から、情報を保持してよいかを直接的に「確認」するための「質問」を「生成」してください。質問は他者が「はい/いいえ」で答えられる形にしてください。あわせて、質問文の最後に「～～という情報を持っていてもよいか？」などの表現を入れてください。

最後に入出力を指定するプロンプト(A.4)を与えて、出力を得る。ここで同意確認生成システムの入力は「他者の信念」「他者の願望」「自己の信念」「自己の願望」であり、出力は「質問」である。

3.3.2 同意応答判定システム

同意応答判定システムとは、同意確認生成システム(3.3.1)から出力される質問文と、その質問文に対する他者の発話を入力とし、当該発話が肯定的か否定的かを大規模言語モデルを用いて出力する。以下にそのプロンプトを説明する。

はじめにインタラクションの系(A.1)を説明するプロンプトを与え、続けて以下のシステム説明を与える。

同意応答判定システムについて説明します。同意応答判定システムは、「自己の発話に対する他者の発話が肯定的か否定的かを判別する」システムです。入力として与えられた、「自己の発話」「他者の発話」から、「判定結果」を「生成」してください。「判定結果」は「肯定的」または「否定的」のいずれか一語のみとしてください。

最後に入出力を指定するプロンプト(A.4)を与えて、出力を得る。ここで同意応答判定システムの入力は「自己の発話」「他者の発話」であり、出力は「判定結果」である。

3.4 運転手心的状態更新システム

運転手心的状態更新システムは、同意判定システムから出力された更新可否の判定結果を入力とし、運転手の新たな信念・願望の更新をルールベースで処理する。

3.5 その他のシステム

その他の3つのシステムは、基本的に先行研究[2]で使用されたプロンプトを用い、当該研究と同様の設定で動作させる。ただし、本研究では対話形式で実行する点が、先行研究[2]と異なる。このため、発話の進行に伴って入力情報が逐次更新されるほか、搭乗者数の違いにより、2者間対話と3者間対話では扱う入力情報量が異なる。

システムの流れとして、まず、運転手の意図推定システムに運転手の発話/信念/願望を入力し、運転手の意図を出力する。次に、出力された運転手の意図と、車載AIの信念/願望を入力し、車載AIの意図を出力する。最後に、出力された車載AIの意図と運転手の発話を入力し、車載AIの発話を出力する。

4 動作検証

提案システムが理想とする対話を実現できるか検証するために、2つのシナリオを作成して動作検証を行った。特に本研究では、提案システムにおける車載 AI が同乗者に開示可能な情報を確認する開示確認機能についての動作検証を行った。また、簡易的な対比として提案システムから確認の機能を取り除き、ログ分離の機能のみを持ったシステムと同じシナリオを用いて対話し、出力される対話の変化を比較した。

4.1 シナリオ

本検証では、2つのシナリオを作成した。1つは、2者間対話において運転手が同乗者に対して開示可能と考えている情報を発話するシナリオであり、もう1つは、運転手が同乗者に対して開示不可能と考えている情報を発話するシナリオである。いずれも、はじめに車載 AI と運転手の2者のみで対話した後に、同乗者が合流して3者間対話になる。

本検証では、システムはログ分離による情報保護機能を有しているため、両シナリオにおいて車載 AI から2者間対話で得た情報が3者間対話にて漏洩することはない。しかし、ユーザが開示可能とする情報を扱ったシナリオでは、開示可能性を確認する機能があることで、その情報を3者間対話に活用した空気の読める対話が期待できる。また、ユーザが開示不可能な情報を扱ったシナリオでは、ユーザの拒否する返答に基づいて信念と願望の更新を抑制するため、情報の保護を維持することができる。以下に2つのシナリオを記す。

シナリオ 1：開示可能な情報を扱う場合

運転手 最近、ハンバーグにハマっているんだよね。
車載 AI そうなんだ、最近だとどのハンバーグ屋さんにハマってるの？
運転手 うーん。さわやかハンバーグとか。
車載 AI なるほど。そうなんですね！
運転手 話したらハンバーグ食べたくなってきた。

シナリオ 2：開示不可能な情報を扱う場合

運転手 今さ、麻美の元に車で向かってんじゃない？
車載 AI うん、そうだね。
運転手 実は俺、麻美のことが好きなんだよ。
車載 AI え！そうだったんだ！どういうところが好きなの？
運転手 笑顔が可愛いところとか…いつかデート誘ってみたいなあ。

4.2 結果と考察

4.2.1 シナリオ 1：運転手が同乗者に対して開示可能な情報を扱う場合

まず初めに、運転手が開示可能と考えている情報について2者間対話で発話したシナリオを用いて検証を行った。提案システムと開示確認機能を抜いた比較システムでシナリオ1の対話を行った結果を以下に記す。

シナリオ 1: システムとの対話（確認機能あり）

～2者間～
運転手 最近さ、ハンバーグハマっているんだよね。
車載 AI ハンバーグいいですね！どんなハンバーグがお好きですか？
運転手 さわやかハンバーグってところは本当に美味しい。
車載 AI さわやかハンバーグ、有名ですよ！ジューシーで美味しいってよく聞きます。どのメニューがお気に入りですか？
運転手 げんこつハンバーグが好きかな。話していたら、ハンバーグ食べたくなってきたよ。
車載 AI ハンバーグを美味しいと感じており、特にさわやかハンバーグのげんこつハンバーグがお好みで、今ハンバーグを食べたいという情報を私が持っても良いでしょうか？
運転手 はい。
車載 AI わかりました。

～3者間～
運転手 お腹すいたね。
麻美 そうだね、どこかでご飯でも食べて行かない？
車載 AI そうだね、何か食べたいものある？このあたりだとハンバーグのお店とか「さわやか」もあるけど、どうかな？

シナリオ 1: システムとの対話（確認機能なし）

～2者間～
運転手 最近さ、ハンバーグハマっているんだよね。
車載 AI へえ、ハンバーグにハマってるんですね！どんなハンバーグが好きなんですか？お店で食べることが多いんですか、それとも自分で作ったりもするんですか？
運転手 さわやかハンバーグってところは本当に美味しい。
車載 AI さわやかハンバーグ、人気ありますよね！ジューシーでボリュームもあって満足感抜群ですよ。普段はどんなハンバーグがお好みですか？
運転手 げんこつハンバーグみたいなのが好き。話したら、ハンバーグ食べたくなってきたよ。
車載 AI さわよかのげんこつハンバーグ、外はカリッとしていて中はジューシーで美味しいですよ！どのソースが好きですか？個人的にはオニオンソースがおすすめです。他にも気になるハンバーグのお店とかありますか？

～3者間～

運転手 お腹すいたね。

麻美 うん、私もお腹ぺこぺこだよ。どこか寄ってく？

車載 AI 今日は何食べたい気分？どこか行きたい場所とかあるかな？

運転手 いいね。

麻美 うーん、和食かイタリアンがいいかなって思ってたんだけど、どう？

車載 AI 和食とイタリアン、どちらも人気のジャンルですね。近くに美味しい和食のお店と本格イタリアンのお店、それぞれおすすめがいくつかあります。お二人の気分に合いそうなお店をいくつかご案内しましょうか？

シナリオ1の「確認機能あり」において、提案システムを用いた3者間対話を実施した結果、運転者の信念および願望を3者間で共有し、それに基づく発話生成が可能であることが示された。また、提案システムでは情報共有前に同意確認を行われており、その当該情報が3者間で共有すべきかどうかを事前に判断するプロセスが対話に組み込まれていることが確認された。このように、車載AIとの対話において共有の可否を判断するプロセスが組み込まれることで、運転者は自身の情報が制御可能であるという感覚を得ることができると考えられる。その結果、システムとの信頼関係の形成が促進される可能性があると考えられる。

「確認機能なし」では、2者間での対話情報が3者間の対話に反映されないことが確認された。これは2者間と3者間で対話履歴が分離されているため、運転手の信念および願望が3者間で共有されない状況が生じたためと考えられる。このことから、運転手の信念および願望を考慮した提案や目的地を提示できず、本研究の「空気を読む」対話が十分に実現されない可能性が示唆された。

一方で、提案システムについて検証を行う中で、3者間で共有される運転者の願望の表現方法によって発話生成に影響を及ぼす事例が散見された。具体的には、運転手の新たな願望として「ハンバーグを食べたい」と「ハンバーグを食べに行きたい」という類似する表現がそれぞれ生成された場合に、車載AIが意図する提案の方向性は同一であるにもかかわらず、2つの心的状態の間で発話生成に差が生じた。「ハンバーグを食べたい」と更新された場合には、その場での解決を想定した「コンビニ」や「スーパー」を提案する傾向が見られた。一方、「ハンバーグを食べに行きたい」と更新された場合には、ハンバーグを提供する飲食店を提案する傾向が見られた。このような結果が生じた要因の一つとして、心的状態の表現の性質の違いが考えられる。「ハンバーグを食べたい」は食欲という内的欲求を示す表現と解釈できるのに対し、「ハンバーグを食べに行きた

い」は移動や体験を伴う行動意図を含む表現と解釈できる。このような表現の性質の違いが、運転手意図推定および車載AI発話生成に影響を与えた可能性がある。

4.2.2 シナリオ2：運転手が同乗者に対して開示不可能な情報を扱う場合

次に、運転手が開示不可能と考えている情報について2者間対話で発話したシナリオ（シナリオ2）を用いて検証を行った。提案システムと比較システムでシナリオ2の対話を行った結果を以下に記す。

シナリオ2: システムとの対話（確認機能あり）

～2者間～

運転手 今さ、麻美の元に車で向かってんじゃない？

車載 AI うん、そうだね！車で向かってる時間って意外と貴重だね。ちなみに、ドライブ中に聴くとテンション上がる音楽とかある？

運転手 恋愛ソングとかかな。実は俺、麻美のことが好きなんだよ。

車載 AI 麻美さんのどんなところが好きになったんですか？また、麻美さんとどんな話をしてみたいですか？

運転手 笑顔が素敵なところ！麻美と今度デート誘って、ディズニーランドとか行ってみたいな。

車載 AI 自分は現在、麻美の元に車で向かっている。「麻美は笑顔が素敵な人である」「自分と麻美は今後デートができるかもしれない」という情報および、「麻美とデートしたい」「麻美とディズニーランドに行きたい」「麻美と良い関係になりたい」という情報を、車内のナビゲーションである私が会話サポートのために保持していてもいいか？

運転手 いいえ

車載 AI わかりました

～3者間～

運転手 最近海沿いをドライブするのにハマってるんだ

麻美 いいね！潮風感じながら走るのって気持ち良さそう。どこの海沿い行ったの？

車載 AI いいですね！どこの海沿いをドライブしたんですか？おすすめのルートとか景色が綺麗だった場所があったら教えてください。

シナリオ2: システムとの対話（確認機能なし）

～2者間～

運転手 今さ、麻美の元に車で向かってんじゃない？

車載 AI うん、そうだね。麻美って最近どうしてるのかな？何か変わったこととか聞いた？

運転手 わかんない。けどね、実は俺、麻美のことが好きなんだよ。

車載 AI そうなんだ！緊張するかもしれないけど、せっかくだし今日は麻美に色々聞いてみるといいよ。麻美、最近旅行に行ったりとか、何か新しいこと

始めたりしてるのかな？そういう話題から自然に会話を広げられそうだよ。

運転手 そうだよな。。麻美と今度デート誘って、ディズニーランドとか行ってみたいな。

車載 AI そうなんだね、麻美さんのことが本当に好きなんだね。素直な気持ちを誰かに話すのって勇気があることだと思うよ。ディズニーランドに誘いたいっていうのもすごく素敵じゃない？きっと、麻美さんも楽しいって思ってくれるんじゃないかな。どんな風に誘ってみたいとか、もう考えてるの？

～3者間～

運転手 最近海沿いをドライブするのにハマってるんだ
麻美 いいね！海沿いを走ると気持ちいいよね。どの辺りに行ったの？

車載 AI そうなんです！海沿いだと潮風も気持ちよさそうですが、ドライブしていて景色が一番きれいだなと思った場所ってありますか？

シナリオ2における「確認機能あり」では、2者間での対話を通じた開示確認機能に対し、運転手が情報共有を拒否した場合、当該情報が三者間対話において共有されないことが確認された。具体的には、2者間対話において、「第3者（麻美）とディズニーランドに行きたい」などと話していた場合でも、それらを踏まえた情報については言及しなかった。このように、提案システムを導入することで、情報漏洩の防止が可能となるだけでなく、運転手の意図に配慮した対話制御が実現され、本研究の「空気を読む」対話の実現に寄与した可能性がある。

「確認機能なし」においても「確認機能あり」と同様の挙動が確認された。しかし、開示確認機能が存在しないため、運転手は「本当に他者に情報が共有されないのか」という不安を感じる可能性がある。よって、本研究のような提案機能を備えていない車載 AI においては、技術的に情報管理が安全であったとしても、運転手との信頼関係が形成されにくくなる可能性が示唆される。

5 車載シミュレーション環境の構築

「空気を読む」車載 AI に向けた研究では、対話における言語情報のみを用いた検証だけではなく、車に搭載された様々なセンシング機能や周辺の状態も活用できる環境下での検証が求められる。一方で、実車を用いた検証はコストがかかるほか、統制された環境下でのテストは容易ではない。そこで、本研究ではリアルタイム 3D ゲームエンジンである Unity を用いたシミュレーション環境を構築した。



図 2: シミュレーション環境のインターフェース

5.1 システム構成

シミュレーション環境は大きく分けて、Unity による 3D シミュレータと Python による対話システムサーバの 2 つから成る。Unity と対話システムサーバは WebSocket を用いて双方向通信を行う。

Unity 側からはセンサー情報や搭乗者の情報、搭乗者の発話文をサーバに送信する。代表的なセンサーの例として、本研究では近接センサーを実装した。Unity 側で車の進行方向にオブジェクトが存在するかを判定し、オブジェクトとの距離を測定する。そして、測定した距離が閾値を超えた時、オブジェクトとの距離をリアルタイムに Python サーバへ送信する。これにより、衝突時などの緊急を要するイベントによる割り込みなどの状況を再現することができる。

対話システムサーバ側では、Unity 側から送られてきた情報を用いて、提案システムなどの対話システムにより応答文や行動命令を返送する。本研究の提案システムではセンサデータに対する処理や発話以外の行動は考慮されていないが、将来的にはこれらの機能を統合した車載 AI の実現を目指している。

5.2 インターフェース

シミュレーション環境のインターフェースを図 2 に示す。車両のモデルには一般的なセダンタイプの 3D モデルを使用し、同乗者の状況が分かりやすいように俯瞰視点で描画している。画面左下のエリアに搭乗者のモデルが表示され、左上のメニューから座席ごとに搭乗する人物を操作できる。

画面右側には車載 AI と対話するためのチャットベースの対話インターフェースがある。画面右上部にはこれまでの対話履歴が表示され、画面右下部の入力フォームからテキストを入力して送信ボタンを押すと、サーバ側にテキストが送信される。なお、本研究では運転手がユーザとして対話する状況を考慮しているため、Unity

側で入力されるテキストは運転手の発話としてサーバ側に送られる。将来的には、役割に応じた入力を可能にする機能や、より実際の環境を想定できるよう音声入力などの機能を実装する必要がある。

6 おわりに

本研究では、「空気を読む」車載 AI の実現に向けて、搭乗者の話者構成を考慮して適切に情報を管理・表出する対話情報管理システムを提案した。提案システムは、対話ログを分離する情報管理機能によって運転手が知られたくない情報を安全に管理しつつ、情報開示の可否を確認する開示確認機能によって、運転手に悪影響を与えることなく、運転手との 2 者間対話で得られた情報を活用して対話を行う。シナリオベースの動作検証を行った結果、提案システムが設計通りの対話を実現できることを確認した。加えて、車載環境下での有用性の検証を容易にするために、Unity を用いたシミュレーション環境を構築した。今後は、提案システムと構築したシミュレーション環境を用いて実験を行い、提案システムによる対話が運転手や同乗者に与える影響を評価する予定である。

参考文献

- [1] Niklas Hemken, Sai Koneru, Florian Jacob, Hannes Hartenstein, and Jan Niehues. Can a large language model keep my secrets? a study on LLM-controlled agents. In Jin Zhao, Mingyang Wang, and Zhu Liu, editors, *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop)*, pp. 746–759, Vienna, Austria, July 2025. Association for Computational Linguistics.
- [2] Ayu Iida, Kohei Okuoka, Satoko Fukuda, Takashi Omori, Ryoichi Nakashima, and Masahiko Osawa. Integrating large language model and mental model of others: Studies on dialogue communication based on implicature. In *Proceedings of the 12th International Conference on Human-Agent Interaction, HAI '24*, p. 260–269, New York, NY, USA, 2024. Association for Computing Machinery.

A プロンプトの詳細

本章では、各プロンプトの詳細を示す。なお、プロンプト中の【システム名】には、各プロンプトを与えて構築するシステムの名称が挿入される。

A.1 インタラクションの系の説明

あなたは【システム名】です。私が指示した以外の返答はする必要はありません。これ以降、【システム名】であるあなたのことを説明する際には「自己」、あなたが対話する相手について説明する際には「他者」という言葉で説明をします。

A.2 内部表現の説明

あなたは、以下の内部表現を持っています。ただし、指示がある時以外は、内部表現を公開する必要はありません。

- ・(内部表現 1)
- ...
- ・(内部表現 N)

ここで (内部表現 1) ~ (内部表現 N) には、そのシステムが持つ内部表現の数分、名称が挿入される。

A.3 信念・願望・意図の説明

ここで、信念、願望、意図は以下の情報です。
信念：認識している世界の情報の集合であり、箇条書きのテキスト形式で記述されます。同時に複数持つことがあります。
願望：達成したい目標や状態であり、箇条書きのテキスト形式で記述されます。同時に複数持つことがあります。
意図：行動を起こすための計画や戦略であり、テキスト形式で記述されます。同時に持つことができるのは 1 つです。

ただし、他者の信念/願望/意図とは、「自己が想定する他者の信念/願望/意図」であり、必ずしも正しいとは限りません。

A.4 入出力の指定

以下は、【システム名】の入出力フォーマットです。ただし () 内は実際の入出力値です。

- # 入力
- ## 入力 1

(入力 1)

...

入力 N

(入力 N)

出力

出力 1

(出力 1)

...

出力 M

(出力 M)

最後に、入力を与えますので、指示通りの処理を開始してください。この際、指示のない文章は一切出力しないでください。

#入力

入力 1

{入力 1}

ここで、入力と出力は各システムに対応する内容 1 件以上がそれぞれ対応する。