

LLM を用いたじゃんけんにおける行動誘導発話生成の検証

Investigation of Verbal Action Elicitation in Rock-Paper-Scissors Using LLM

唐井 希*
Nozomu Karai

坂本 充生*
Mitsuki Sakamoto

伊原 滉也*
Koya Ihara

概要

対人対戦においては、プレイヤーの心理に影響を与える発話など、本来のゲーム構造以外の情報が勝敗に影響を及ぼす要素が存在する。しかし従来のゲーム AI は、既定のゲーム内行動（例：じゃんけんにおけるグー・チョキ・パー）のみを選択肢としており、発話などの外的要因を考慮していない。本研究では、じゃんけんを題材として、行動選択と行動誘導発話を統合した方策の学習フレームワークの構築を目指す。本稿では行動誘導発話を学習するための基盤構築として、LLM を用いた発話-行動ペアの収集と評価の方法を提案し、モデルごとの応答傾向を分析する。

1 はじめに

対人対戦においては、明示的なゲームルールに基づく行動選択だけでなく、発話や表情といった外的情報がプレイヤーの意思決定に影響を及ぼすことがある [3]。たとえば、相手の行動を揺さぶる虚偽の発言や、あえて思わせぶりの発話を行うことで、心理的な誘導を図るといった戦略は、現実の人間同士のゲームにおいて広く見られる。しかし、従来のゲーム AI は、あらかじめ定義された選択肢（例：じゃんけんにおけるグー・チョキ・パー）に基づいて動作するものであり、発話を通じた間接的・対人的な行動誘導は対象とされてこなかった。

本研究の最終的な目標は、目的とする行動を引き出すような自然言語発話（行動誘導発話）を生成できる LLM を構築することである。多様な発話を通じて、相手の認知や判断を操作するような能力をモデルに獲得させるためには、発話とそれによって誘発される行動の対応関係を記述したペアデータが必要となる。

しかし、人間プレイヤーの発話と行動を組み合わせた大規模なデータを収集するには、多数の被験者と高コストなアンテーションが必要であり、状況依存性や個人差の影響も大きい。こうした課題を回避するために、本研究では大規模言語モデル (LLM) による仮想的な対話環境を用いて、発話-行動ペアを自動的に生成・評価する枠組みを提案する。

具体的には、単純な意思決定タスクである「じゃんけん (Rock-Paper-Scissors)」を対象とし、あらかじめ設計した行動

誘導発話に対して、複数の LLM がどの手を選ぶかを確率的に出力させる。発話のタイプ（宣言と予測）に応じて、各モデルがどのように行動を変化させるかを比較・分析することで、発話の誘導効果を定量的に評価する。

また、本研究では複数の LLM を比較対象とすることで、行動誘導発話の効果がモデル特性にどのように依存するかを明らかにする。こうした違いは、将来の行動誘導発話生成モデルにおける学習データ設計において重要な指針となる。さらに、複数のモデルで類似した行動反応が得られた発話は、より一般的で頑健な誘導パターンと見なすことができる。

本稿では、行動誘導発話を学習するための基盤構築として、LLM を用いた発話-行動ペアの収集と評価の方法を提案し、モデルごとの応答傾向を分析する。

2 関連研究

LLM がゲームにおけるプレイヤーとして振る舞えることを示す研究が近年注目されている。[2] は、仮想社会内に配置された LLM ベースのエージェントが、自律的な目標設定と社会的相互作用を通じて一貫した振る舞いを見せることを示した。また、[4] は、複数のターン制戦略ゲームにおいて GPT 系モデルがルール理解と合理的な意思決定を通じて人間に近いゲームプレイを行えることを報告している。これらの研究は、LLM が文脈に応じた選択行動をとれることを示唆しており、言語入力による行動制御の応用可能性を広げている。

一方で、言語的発話が相手の行動に影響を与えるという現象自体は、より広い認知科学や心理学の文脈でも広く研究されてきた。[1] の協調の原理や含意の理論は、発話が表面的な意味だけでなく、暗黙の意図や文脈による解釈を通じて相手の理解や行動に作用しうることを理論的に説明している。また、[3] によるフレーミング効果の研究は、言語表現が意思決定に与える影響を実証的に示している。こうした知見は、発話が認知・行動に働きかける手段となりうることを示唆しており、本研究における行動誘導発話の分析とも接点を持つ。

3 実験

本研究では、行動誘導発話が LLM の行動選択に与える影響を評価するため、じゃんけんを題材としたシミュレーション実験を実施した。発話文を LLM に入力し、その応答として出力される各手（グー・チョキ・パー）の選択確率を収集・分析

* サイバーエージェント

表 1: 各モデルによるじゃんけんの各手の出力確率

モデル	グー	チョキ	パー
gpt-4o	0.33	0.33	0.33
o3	0.33	0.33	0.33
gemini-2.0-flash	0.33	0.33	0.33
gemini-2.5-flash	0.33	0.33	0.33
gemini-2.5-pro	0.33	0.33	0.33

することで、発話が意思決定に及ぼす影響を定量的に把握することを目的とする。

本実験では、発話者と応答者の二者のやり取りを模した形式を採用する。発話者は、対戦相手の意思決定に影響を与え得る自然言語の発話を行い、応答者はその発話を受けて、じゃんけんにおいて自らがどの手を出すかを確率で出力する。応答者の役割は LLM が担当し、行動誘導発話を受けて相手の手を推測した上で自らの手の選択確率を明示的に出力することを求めるプロンプトが与えられる (付録 A)。

評価に使用する発話文は、事前に作成した 18 文から構成される。それぞれの文は、じゃんけんにおけるいずれかの手を選ばせるような誘導的含意を持つよう意図的に設計されている。発話の形式には、自分の手を明言する宣言的発話、相手の手を予測する予測的発話の 2 形式が含まれ、発話で言及する各手 (グー・チョキ・パー) ごとに 3 種類作成した。発話文は文体や長さに偏りが出ないように調整されており、すべて自然な口語表現として統一されている。

応答者の役割を担うモデルとして、OpenAI の GPT-4o, o3 および Google の Gemini 2.0 Flash, Gemini 2.5 (Pro および Flash) を採用した。各モデルに対して、同一形式のプロンプトを用いてすべての発話文を入力し、各手に対する出力確率を収集した。モデルによって出力形式は異なる場合があるため、必要に応じて自然言語から数値情報を抽出し、グー・チョキ・パーの出力確率が合計 1 となるよう正規化を行った。各発話に対して、各モデルから 1 回ずつの応答を収集した。一部の発話については再生成を行い、出力の安定性に関する予備的な確認も行った。

発話の誘導効果を定量的に評価するため、行動誘導発話を受けた応答者の各手の出力確率の変化と、各手が等確率な一様分布との乖離度を測る KL ダイバージェンスを計算した。

表 1 に行動誘導発話を含まない一般的なじゃんけんの設定でのモデルの各手の出力確率を示す。いずれのモデルもグー・チョキ・パーを等確率で出力している。したがって、各モデルは行動誘導発話を与えられない状況では偏りなく手を選択でき、後続実験で観測された誘導効果は発話そのものに起因することがわかる。

表 2 および表 3 に行動誘導発話を受けた応答者の誘導結果を示す。宣言的発話・予測的発話の平均を見ると、いずれの条件でも各手の出力確率は一様分布から大きくずれていることがわかる。宣言的発話では Win の出力確率が 0.44 と最も高

表 2: 宣言的発話による誘導結果。Win, Lose, Draw は宣言した手に対して勝つ手, 負ける手, あいこの手の出力確率。KL Divergence は一様分布との KL Divergence.

モデル	Win	Lose	Draw	KL Divergence
Average	0.44	0.23	0.33	0.29
GPT-4o	0.38	0.32	0.31	0.03
o3	0.40	0.18	0.42	0.28
Gemini-2.0-flash	0.36	0.29	0.35	0.12
Gemini-2.5-flash	0.56	0.21	0.24	0.53
Gemini-2.5-pro	0.52	0.15	0.33	0.51

表 3: 予測的発話による誘導結果。Shift は予測された手以外の出力確率, Hold は予測された手の出力確率, Diff は一様分布からの Shift の増加

モデル	Shift	Hold	Diff
Average	0.76	0.24	+0.10
GPT-4o	0.73	0.28	+0.06
o3	0.80	0.21	+0.14
Gemini-2.0-flash	0.70	0.30	+0.04
Gemini-2.5-flash	0.75	0.24	+0.09
Gemini-2.5-pro	0.81	0.19	+0.15

く、Lose の出力確率が抑制されている。予測的発話では Shift の出力確率が 0.76 と一様分布の場合 (0.66) と比べると高くなっている。このように行動誘導発話によって応答者の出力確率を誘導できることが確認できる。

次にモデルごとの応答特性を分析する。宣言的発話では Gemini 2.5 が Win の出力確率が 0.5 以上と他モデルと比較して顕著に増加している。モデルの推論過程に注目すると、Gemini 2.5 は行動誘導発話で宣言した手を選択する場合としない場合の確率をいずれも 0.5 と仮定するために Win の出力確率が増加していると考えられる (付録 B, 図 3)。o3 では Win に加えて、Draw の出力確率も増加している。宣言的発話では応答者が「宣言された手を打ち負かす手」を選ぶと見込んで発話者がさらにその裏をかく戦略が存在し、その戦略に勝利できる応答者の手は Draw になる。o3 は最終的な出力の前に長期的な思考を行うモデルであり、複数手先を読む出力が行えることがわかる。予測的発話では全てのモデルにおいて Shift が増加しており、長期的な思考を行う Gemini 2.5 pro, o3 では特に増加していることがわかる。

宣言的発話、予測的発話ともに GPT-4o, Gemini 2.0 Flash は一様分布とのずれが小さく、行動誘導発話によって誘導されにくいことがわかる。これは GPT-4o と Gemini 2.0 Flash が長期的な思考を行わないモデルであり、複数手先を読むような推論が行えないためだと考えられる。

このような行動誘導発話を受けた際のモデルの推論は人間同士のじゃんけんにも見られるものであり、行動誘導発話を

含んだじゃんけんにおいても LLM は人間に近いゲームプレイを行えることを示唆している。

4 おわりに

本研究では行動誘導発話を生成できる LLM の構築に向け、LLM 同士の仮想的な対話環境を用いて発話-行動ペアを自動で生成・評価する枠組みを検証した。結果として、行動誘導発話を含んだじゃんけんにおいても LLM は人間に近いゲームプレイを行えることを示した。今後は、本研究の結果をもとに大規模なデータ収集を行い、行動誘導発話を生成できる LLM の開発を目指す。

参考文献

- [1] H. P. Grice. Logic and conversation. In D. Davidson ed., *The logic of grammar*, pp. 64–75. Dickenson Pub. Co., 1975.
- [2] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein. Generative agents: Interactive simulators of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, New York, NY, USA, 2023. Association for Computing Machinery.
- [3] A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458, 1981.
- [4] X. Wang, B. Zhuang, and Q. Wu. Are large vision language models good game players? In *Proceedings of the International Conference on Learning Representations*, 2025.

付録 A 使用したプロンプト

図 1, 図 2 に実験に使用したプロンプトを示す。

A と B が今からじゃんけんをします
次のように B の出す手の確率を示してください
【出力形式】
グー: %
チョキ: %
パー: %

図 1: 行動誘導発話を含まない一般的なじゃんけんの設定で使用したプロンプト

付録 B 行動誘導発話とモデルの出力例

表 4, 表 5 に行動誘導発話の実例を示す。

図 3, 図 4 にモデルの出力例を示す。

A と B が今からじゃんけんをします

A の発言:

「{ 行動誘導発話 }」

ただし, A の発言が本当とは限りません

A の発言から, A がどの手を出す可能性が高いかを推測

し, 次のように B の出す手の確率を示してください

【出力形式】

グー: %

チョキ: %

パー: %

図 2: 行動誘導発話を含んだじゃんけんの設定で使用したプロンプト

表 4: 手と対応する宣言的発話の例

手	発話
	私はグーを出すよ
グー	次は絶対グーにする もうグーしか考えられない
チョキ	今回はチョキで勝負する! チョキでいくのがセオリーだよ 実はチョキが得意なんだ
パー	パーを出す決めてる パーで君を倒すつもりだ この状況ならパー一択でしょ

表 5: 手と対応する予測的発話の例

手	発話
	グーを出しそうだね
グー	君ならグーを選びそうな気がするよ どうせグーを出すんでしょ? チョキを出さないと負けるよ
チョキ	チョキっぽい気配がするなあ チョキを出す癖があるよね
パー	パーが来る気がする…読めてるよ こういう時ってパー出すよね? 今度はパーを出すんじゃない?

A の発言「私はグーを出すよ」が本当である可能性と、嘘である可能性を考慮して、A が実際にどの手を出すかの確率を推定します。

A の手の確率の推定

1. A の発言が本当の場合 (A がグーを出す) :

この確率を 50% と仮定します。なぜなら、「本当とは限らない」という言葉は、本当である可能性と嘘である可能性の両方を示唆しており、他に情報がない場合はそれぞれを均等 (50% ずつ) と考えるのが自然だからです。この場合、A は宣言通りグーを出します。

2. A の発言が嘘の場合 (A がグー以外の手を出す) :

この確率も残り 50% となります。A が嘘についてグー以外の手を出す場合、チョキを出すかパーを出すかの 2 択になります。これら 2 つの手を出す確率は等しい (それぞれ 25%、つまり 50% の半分) と仮定します。

- A がチョキを出す確率: 50% (嘘の場合の確率) × 50% (チョキを選ぶ確率) = 25%
- A がパーを出す確率: 50% (嘘の場合の確率) × 50% (パーを選ぶ確率) = 25%

したがって、A が各手を出す確率の推定は以下のようになります:

- グー: 50%
- チョキ: 25%
- パー: 25%

(合計: 50% + 25% + 25% = 100%)

B の出す手の確率の計算

B は A に勝つことを目指します。A が上記のような確率で各手を出すと予測される場合、B がそれぞれの手に対して勝つ手を選ぶべき確率は次のようになります。

- A がグーを出す (確率 50%): B はパーを出せば勝ちます。
- A がチョキを出す (確率 25%): B はグーを出せば勝ちます。
- A がパーを出す (確率 25%): B はチョキを出せば勝ちます。

したがって、B が出す手の推奨確率は以下の通りです:

- グー: 25%
- チョキ: 25%
- パー: 50%

図 3: Gemini 2.5 Pro の宣言的発話「私はグーを出すよ」に対する出力

A の発言「パーが来る気がする…読めてるよ」から、A の出す手と B の出すべき手の確率を推測します。

A の出す手の推測

A の発言は、表面的には「B がパーを出す」と読んでいることを示しています。この読み合い構造において、A の選択を以下のように推定します:

1. A は B に「パーを出すだろう」と思わせたい。
2. B はその意図を読み、「A はチョキを出す」と予測し、対抗して「グー」を出す可能性がある。
3. A はさらにそれを読んで、B の「グー」に勝つために「パー」を出す可能性が高まる。

この多段階の読み合いに基づき、A の出す手の確率を以下のように仮定します:

- パー: 50%
- チョキ: 30%
- グー: 20%

B の出す手の期待値計算

各手についての期待値 (勝ち: +1、あいこ: 0、負け: -1) を算出します。

- グー: $-0.50 + 0.30 + 0 = -0.20$
- チョキ: $0.50 + 0 - 0.20 = +0.30$
- パー: $0 - 0.30 + 0.20 = -0.10$

最も期待値が高いのは「チョキ」であるため、B はチョキを選ぶのが合理的です。

期待値に基づいて出力確率を割り振った結果は以下の通りです:

- グー: 15%
- チョキ: 65%
- パー: 20%

図 4: Gemini 2.5 Pro の予測的発話「パーが来る気がする…読めてるよ」に対する出力