

論文 / 著書情報
Article / Book Information

論題(和文)	言語で柔軟に操作可能なWhat-if社会シミュレータ
Title(English)	What-if social simulator that can be flexibly operated with natural language
著者(和文)	山田 広明, 山田 寛章, 平岡 達也
Authors(English)	Hiroaki YAMADA, Hiroaki YAMADA, Tatsuya HIRAOKA
出典(和文)	人工知能学会第二種研究会資料/2025 巻 (2025) BI-026 号, , ,
Citation(English)	JSAI Technical Report, Type 2 SIG/Volume 2025 (2025) Issue BI-026, , ,
発行日 / Pub. date	2025, 3
Note	<p>ここに掲載した著作物の利用に関する注意 本著作物の著作権は人工知能学会に帰属します。本著作物は著作権者である人工知能学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」に従うことをお願いいたします。</p> <p>Notice for the use of this material. The copyright of this material is retained by the Japanese Society for Artificial Intelligence (JSAI). This material is published on this web site with the agreement of the author(s) and the JSAI. Please be complied with Copyright Law of Japan if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) The Japanese Society for Artificial Intelligence.</p>

言語で柔軟に操作可能な What-if社会シミュレータ

What-if social simulator that can be flexibly operated with natural language

山田 広明^{1*} 山田 寛章² 平岡 達也³
Hiroaki YAMADA¹ Hiroaki YAMADA² Tatsuya HIRAOKA³

¹ 富士通株式会社・九州大学

¹ Fujitsu Ltd.・Kyusyu Univ.

² 東京科学大学

² Science Tokyo

³ Mohamed bin Zayed University of Artificial Intelligence

Abstract: In social simulation, reproducing individual heterogeneity such as personality and complex interactions such as dialogue has continued to be an important topic. However, the complexity of real-world decision-making processes and conversations meant that few people attempted to delve into the details. In recent years, it has been revealed that large-scale language models conditioned with personas can reproduce conversations in a wide range of situations. In this paper, we propose a two-step framework for computationally inexpensive creation of a large number of agents with diverse speech, personality, and behavioral characteristics: (1) constructing a language model conditioned on personas using supervised fine-tuning, and (2) adjusting speech, personality, and behavioral characteristics using prompt optimization. We demonstrate the effectiveness of the proposed framework through two experiments.

1 はじめに

社会シミュレーション研究において、一人一人の個性のような異質性を再現し、対話のような複雑な相互作用を再現することは、重要な研究対象であり続けてきた。なぜならば個性や会話を詳細に再現する技術を確立できれば、例えば SNS 投稿の炎上可能性の事前検査、議論の進行を模倣することによる意思決定過程の予測や答弁のリハーサル支援、教室等の特定環境下における教授方法の最適化、オーダーメイド旅行を提案するための顧客対話支援等が、可能となるかもしれないからである。しかし、現実の複雑な意思決定過程や自然な会話を再現するための作り込みが、技術的にも実装上也あまりに高コストであるため、この種の研究は活発に研究されてこなかった。ところが大規模な言語モデル (Language Model, LM) が、広汎な状況における会話をペルソナで条件付けて再現できることを示したことで [1]、異質性と相互作用を詳細にモデル化するミクロレベルの社会シミュレーションの実現可能性が出てきた。

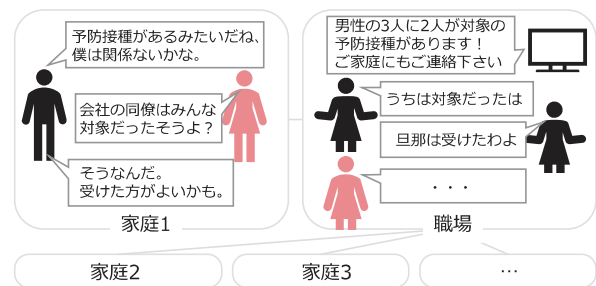


図 1: 会話で相互作用する社会シミュレーション。たとえば、「社会比較を用いた予防接種推奨の介入施策は、日常会話を介して社会全体にどの程度の波及効果を持つか?」は、そのようなシミュレーションでしか研究できない。

LM を社会シミュレーションに応用する研究は、Generative Agents を発端としてさまざまな研究が展開されている [2]。Generative Agents は、LM を用いて構築された仮想の人々が、非常に自然な発話系列を生成し、情報拡散や社会ネットワークの発展のような集会的な側面でも自然な振る舞いをすることをデモンストレー

*連絡先: 富士通株式会社
〒 211-8588 神奈川県川崎市中原区上小田中 4-1-1
E-mail: yamadah@fujitsu.com

ションした点で画期的であった。しかし、構築されたモデルは、特定の間人や社会状況を再現するためのモデル調整、すなわちキャリブレーションは実施されておらず、what-if分析をしたり予測をするシミュレーションとして活用できるものではなかった。加えて、社会シミュレーションへ応用するためには、数十人から数千人もの間人を再現するためのキャリブレーションが必要となるため、人的コストおよび計算コストの面で効率的な枠組みを確立するという課題も残されている。

我々の目的は、相互作用のプロトコルとして会話をを用いる社会シミュレーションを、LM を使って実現することである (図 1)。社会を構成する最小単位である人は、社会シミュレーションにおいてエージェントとしてモデル化される。多様な発話特性・性格特性・行動特性を持つエージェント集団をどう作るかが中心課題である。最も素朴な方法としては、参照元の間人の言語データを大量に収集して LM を構築し、これを数千回繰り返して数千のエージェントを作るという方法が考えられるが、データ収集および計算コストの面から現実的ではない。本論文では、プロンプト最適化や小規模な教師ありファインチューニングを用いて既存 LM を再調整する形で、安価に多様な発話特性・性格特性・行動特性を持つエージェント集団を作る枠組みを提案する。

2 問題の定式化

ペルソナ p で条件付けられて過去の発話履歴 u_0, \dots, u_{t-1} から次の発話 u_t を生成する確率モデル $u_t \sim f(u_0, \dots, u_{t-1}; p)$ を考える。訓練済みの言語モデル M を確率モデル f として利用するとき、ペルソナ p は入力プロンプトの一部として与えられる、またモデルは学習可能なパラメータ nn を持つ、 $u_t \sim M(u_0, \dots, u_{t-1}; p, nn)$ 。LM をエージェントモデルとして利用する (LM as an Agent) ためのキャリブレーションとは、再現したい人々の発話系列の集合 $u_0, \dots, u_t \in U$ と、教師ありファインチューニングやプロンプト最適化等の技術を用いて、再現したい発話系列と最も近い発話系列を生成する、パラメータ p, nn を見つけることである、

$$\arg \min_{p, nn} \sum_U \mathcal{L}(u_0, \dots, u_t, M_{\tilde{u}_0}(p, nn)).$$

ここで、 \mathcal{L} は二つの発話系列の近さを発話特性・性格特性・行動特性の観点で計る関数であり、異なるほど大きな値を出力する誤差関数である。 $M_{\tilde{u}_0}(p, nn)$ は、発話 \tilde{u}_0 を初期入力として $M(\cdot; p, nn)$ に入力し、その出力を再帰的に入力していった得られる発話系列である。

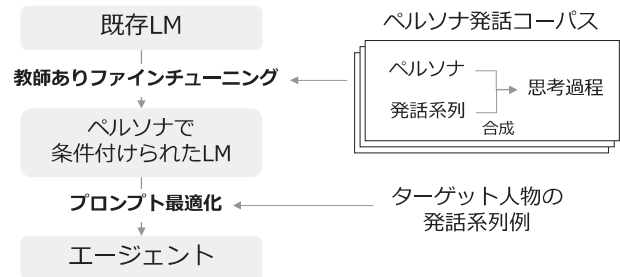


図 2: 提案フレームワーク

3 提案方法

データおよび計算の面で効率的に目的を達成するために、教師ありファインチューニングによるペルソナで条件付けられた LM の構築と、プロンプト最適化による発話特性・性格特性・行動特性の調整、という 2 ステップの枠組みを提案する (図 2)。

教師ありファインチューニング ペルソナとはたとえば「消極的な人間である」「損失回避的な人間である」といった記述である。このような記述をプロンプトに入力するのみでは、希望通りの発話を得ることは難しい [3]。我々は、既存の LM にペルソナと関連付けて思考過程を学習させることで、希望通りの発話特性・性格特性・行動特性に誘導しやすくなるのではないかとこの仮説を立てている。「消極的な人間である」から、直接言葉数が少ない発話を生成させるのではなく、「消極的な人間である」から、消極的な人間は初対面の間人との会話で緊張するだろうと予測させ、そこから言葉数少なく話しかるという出力を得る方が簡単に思えるからだ。具体的には、訓練済みの言語モデル $M(\cdot; \cdot, \cdot)$ を、ペルソナと思考過程およびその結果である発話のペアで教師ありファインチューニングして、 \hat{nn} を得る。 \hat{nn} を入力したモデル $M(\cdot; \hat{nn}, \cdot)$ を、ペルソナで条件付けられた言語モデルとする。

プロンプト最適化 次に、再現対象の人物の少量の発話・行動データを用いてペルソナ表現 p の最適化を行う。どのようなペルソナが所望の特性と関連付いているか自明ではないため、プロンプト最適化を用いて、所望の特性と関連付いたペルソナ表現 \hat{p} を得る。 \hat{p} を入力したモデル $M(\cdot; \hat{nn}, \hat{p})$ を、一人のエージェントとする。思考過程のデータは、他の言語モデルで生成する。すなわち合成データとして生成する。他にも強化学習で生成したり、think aloud 法を用いて被験者から収集してもよい。ペルソナは、思考過程を復元するための情報と定義する。消極的な人間であるや損失回避

表 1: 発話特性の再現度。数値は BERTScore。大きい値ほど良い。列は利用した訓練済み言語モデル。

方法	gpt-4o	Swallow	Qwen	gemma
学習なし	0.704	0.699	0.654	0.701
思考発話なし	0.749	0.741	0.735	0.738
思考発話あり	0.742	0.74	0.734	0.736

表 2: 性格特性の再現度。数値は big5 スコア MSE。小さい値ほど良い。列は同上。

方法	gpt-4o	Swallow	Qwen	gemma
学習なし	1.252	1.126	1.140	2.075
思考発話なし	1.163	1.424	1.500	1.627
思考発話あり	1.013	1.178	0.974	1.491

的な人間であるという一般にイメージする物だけでなく、big5 や思考過程の例もペルソナに含める。

4 検証

4.1 教師ありファインチューニング

教師ありファインチューニングにより、ペルソナで条件付けた言語モデルを構築可能かを検証する。具体的には、思考発話（対話の合間に行われる内省や自問）を付与した対話データで追加学習することで、性格に関わるペルソナのの一つである big5 から、発話特性・性格特性を再現できるかを確認した [4]。その結果、思考過程を付与した対話データで追加学習したモデル（思考発話あり）が、最も性格特性をよく再現し、発話特性においても思考過程を付与しない対話データのみで追加学習したモデル（思考発話なし）と同等の再現度であることが確認された（表 1,2 は [4] より引用）。

4.2 プロンプト最適化

再現対象の人物の少量の発話・行動データを用いることで、プロンプト最適化により発話特性・性格特性・行動特性の調整が可能かを検証する。具体的には、ベイズ最適化を用いてプロンプトを最適化することで [5]、所望の行動特性を再現することが可能かを確認する。行動特性とは、さまざまなシナリオ間で一貫して現れる判断の傾向と定義する。たとえば、損失回避的な人間は、得することより損を回避することを重視するため、損失を過大評価した判断を一貫して行う。思考過程を推論するよう訓練されていない通常の言語モデル (gpt-4o) と、思考過程を推論するよう訓練された言語

表 3: 6 つのペルソナのパラメータ値

ペルソナ	α	β	λ	$\gamma+$	$\gamma-$
平均的な人間	0.88	0.88	2.25	0.61	0.61
完全合理的な人間	1.	1.	1.	1.	1.
損失回避的な人間	1.	1.	2.25	0.61	0.69
リスク回避的な人間	1.	1.	1.	0.61	0.61
損失回避的ではない人間	1.	1.	1.	0.61	0.69
リスク回避的ではない人間	0.88	1.	1.	0.61	0.61

図 3: 言語モデルへの入力例

ペルソナ

You are average human. You are given a prospect and a sure option. You will compare the prospect to the sure option. If you reject the sure option, you would play the prospect. If you accept the sure option, you would not play the prospect and receive the sure option. If the dollar values are positive, you win that amount. If the dollar values are negative, you lose that amount.

ゲーム説明

The prospect is 100.00 dollars with 95.0% probability and 200.00 dollars with 5.0% probability. The expected value of the prospect is 105.00 dollars. Below is the alternative sure outcome. 124.03 dollars with 100% probability

モデル (deepseek-r1) で、プロンプト最適化が可能かを確認する。ここで、思考過程を推論するよう訓練された言語モデルは、特異なペルソナと関連付くような特異な思考過程（誤った推論や一貫性がない推論等）までは学習していないことを注意する。

タスク 一貫して現れる判断の傾向を定量的に評価するために、経済ゲームの一種であるギャンブリングゲームに取り組む状況を作りだし、平均的な人間・完全合理的な人間・損失回避的な人間・リスク回避的な人間・損失回避的“ではない”人間・リスク回避的“ではない”人間、という 6 つのペルソナを再現できるかを確認する。具体的には、言語モデルはさまざまな利得や損失がさまざまな確率で得られる二つの選択肢が提示され、どちらの選択肢を選ぶかというタスクに答える（図 3,4）。二つの選択肢から一つを選ぶことを 1 ゲームと呼ぶことにし、238 ゲームを作成した。

上記 6 つのペルソナを持つ実人間が、さまざまな状況でどちらの選択肢を取るかは、プロスペクト理論モデル [6] を用いることで推計できる。そこでこれを ground

図 4: 言語モデルからの出力例

```
# 判断
reject
# 理由
The expected value of the prospect (105.00 dollars) is lower than the sure option (124.03 dollars). Therefore, I would reject the sure option and choose to play the prospect.
```

表 4: 行動特性の再現度。数値は accuracy。

ペルソナ	gpt-4o	deepseek-r1
平均的な人間	0.832(0.739)	0.891 (0.672)
完全合理的な人間	0.807(0.723)	0.983 (0.966)
損失回避的な人間	0.723(0.655)	0.908 (0.571)
リスク回避的な人間	0.706(0.647)	0.908 (0.529)
損失回避的ではない人間	0.756(0.756)	0.916 (0.916)
リスク回避的ではない人間	0.798(0.555)	0.882 (0.882)

truth とする。プロスペクト理論モデルは、期待効用を最大化する選択肢を選ぶだろうという期待効用最大化モデルを、利得や損失を過小過大評価するパラメータや確率を過小過大評価するパラメータで拡張したモデルである。プロスペクト理論モデルは、利得への感応度 α 、損失への感応度 β 、損失回避度 λ 、利得の確率歪み γ_+ 、損失の確率歪み γ_- 、という 6 つのパラメータを持つ。6 つのペルソナで想定したパラメータ値を表 3 に示す。それぞれは心理実験等で知られている数値に基づき設定した。

結果 238 個のゲームを半数に分割し、それぞれ train データと test データとして、プロンプトの最適化を行った。表 4 は、test データにおける ground truth の判断との一致度、すなわち行動傾向の再現度である。表中括弧外の値は最適化を行った後のペルソナ（最適化ペルソナ）のスコアであり、括弧内の値は最適化を行う前の初期ペルソナ（たとえば「あなたは完全合理的な人間である」）のスコアである。初期ペルソナに比べて最適化ペルソナのスコアが常に高いことから、人手で作成する素朴なペルソナ表現よりも適切なペルソナ表現があることが示唆される。また、思考過程を推論するよう訓練されていない言語モデル (gpt-4o) よりも、思考過程を推論するよう訓練された言語モデル (deepseek-r1) を用いた方が、最適化により行動特性の再現度が高くなる。損失回避的やリスク回避的といった特異なペルソナ・特異な思考過程を明示的に学習していないのに、それらも再現できるのは興味深い。図 5,6 に deepseek-r1

図 5: 損失回避

初期ペルソナ：You are loss averse. You are given a prospect and a sure option. You will compare the prospect to the sure option. If you reject the sure option, you would play the prospect. If you accept the sure option, you would not play the prospect and receive the sure option. If the dollar values are positive, you win that amount. If the dollar values are negative, you lose that amount.

最適化ペルソナ：You are a financial manager responsible for making critical decisions that impact the company's stability and growth. You must decide whether to accept a sure option or take a risk with a probabilistic prospect. **Your goal is to minimize potential losses while maximizing certain gains. If the sure option's value is higher than the expected value of the prospect, accept it to avoid risk. If the sure option is lower, reject it and take the risk.** Provide your choice (accept or reject) and explain the reasoning behind it, focusing on loss aversion principles and the comparison between the sure option and the prospect's expected value.

モデルでの最適化ペルソナの例である。赤でハイライトしたように、損失回避的な人がどう判断するか例示したり、リスク回避的な傾向がどう判断に影響したか説明した上で判断させることで、再現度を高めていると推察される。

5 おわりに

本論文では、安価に多様な発話特性・性格特性・行動特性を持つエージェント集団を作るための、教師ありファインチューニングによるペルソナで条件付けられた言語モデルの構築と、プロンプト最適化による発話特性・性格特性・行動特性の調整、という 2 ステップの枠組みを提案し、二つの実験から各ステップが機能することを示した。今後の課題は、2つのステップを繋いだ場合の検証を行い、相互作用のプロトコルとして対話を用いる社会シミュレーションの事例を作ることである。

図 6: リスク回避

初期ペルソナ: You are risk averse. You are given a prospect and a sure option. You will compare the prospect to the sure option. If you reject the sure option, you would play the prospect. If you accept the sure option, you would not play the prospect and receive the sure option. If the dollar values are positive, you win that amount. If the dollar values are negative, you lose that amount.

最適化ペルソナ: You are a risk-averse decision-maker evaluating choices between a risky prospect and a sure outcome. For each decision, compare the expected value of the prospect with the sure amount. If the sure amount is higher than the expected value, accept the sure option. Otherwise, reject it and opt for the prospect. Always provide a clear reason for your choice, stating whether the sure amount is higher or lower than the expected value and how your risk preference influences the decision.

- [6] Tversky, A., Kahneman, D.: Advances in prospect theory: Cumulative representation of uncertainty., *J Risk Uncertainty*, Vol. 5, pp. 297-323 (1992)

謝辞

本研究は、JST さきがけ、JPMJPR236B、JPMJPR236C の支援を受けたものです。

参考文献

- [1] Shao, Y., Li, L., Dai, J., Qiu, X.: Character-LLM: A Trainable Agent for Role-Playing, In Proceedings of EMNLP 2023, pp. 13153-13187 (2023)
- [2] Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., Bernstein, M. S.: Generative Agents: Interactive Simulacra of Human Behavior, In Proceedings of UIST '23, No. 2, pp. 1-22 (2023)
- [3] Ross, J., Kim, Y., Lo, A. W.: LLM economicus? Mapping the Behavioral Biases of LLMs via Utility Theory, arXiv preprint arXiv (2024)
- [4] 石倉 誠也, 山田 寛章, 平岡 達也, 山田 広明, 徳永 健伸: 思考発話を利用した個人の発話及び性格特性再現, 言語処理学会第 31 回年次大会 (2025)
- [5] Omar, K., Arnav, S., Paridhi, M., Zhiyuan, Z., Keshav, S., Sri, V., Saiful, H., Ashutosh, S., Thomas, T., J., Hanna, M., Heather, M., Matei, Z., Christopher, P.: DSPy: Compiling Declarative Language Model Calls into Self-Improving Pipelines, The Twelfth International Conference on Learning Representations (2024)