

IBL エージェントを用いたローカルエコーチェンバー形成メカニズム とフェイクニュース拡散のシミュレーション

Simulation of the local echo chamber formation and fake news diffusion using IBL agents

佐々木 健矢[†], 森田 純哉[†]

Kenya Sasaki, Junya Morita

[†] 静岡大学

Shizuoka University

sasaki.kenya.21@shizuoka.ac.jp, j-morita@inf.shizuoka.ac.jp

概要

近年では SNS (Social Networking Service) 上におけるフェイクニュースの拡散による意見の分極化などが問題視されている。その原因の一つとして、小さな閉じた社会での意見の偏りが挙げられる。本研究では、これをローカルエコーチェンバーと呼び、その性質を理解するため、IBLT (Instance Base Learning Theory) に基づいて、ローカルエコーチェンバーの形成と、フェイクニュース拡散のシミュレーションを行う。

キーワード: 認知モデル, マルチエージェントシミュレーション, 事例ベース推論

1. 背景・目的

SNS 上で自身と似た意見を持つユーザーと繋がることで、自身の発信と似た意見や情報のみ返ってくる状況であるエコーチェンバー(総務省, 2019)の存在が指摘されている。エコーチェンバーの性質として、森田(2023)は、閉じた小規模の集団内での過信に基づく意見の偏りが生じることに注目し、この側面を強調したローカルエコーチェンバーという用語を導入した。ローカルエコーチェンバーは極化を導き、社会全体として見たときに、流通する情報の信頼を脅かすと考えられる。

この考えは、人間が進化的に獲得されたとされる「開かれた警戒メカニズム」を用いて、得られた情報を評価するという Mercier (2020) の考えと一致する。このメカニズムは、社会における個人は、自身の立場を強化する情報に対する過信(開かれた性質)と自身の信念に反する情報への不信(警戒)を併せ持つとする。開かれた警戒心の存在により、ローカルエコーチェンバーが進展すると同時に、フェイクニュースの拡散も導かれる。人間の開かれた性質は、自身の直感に合う情報への過信を導き、誤情報の拡散につながる (Vosoughi et al. 2018)。

これらを踏まえた上で、本研究ではローカルエコーチェンバーの存在を確かめ、その性質を理解するシミュレーションを実施する。その特徴は、次の3点にまとめられる。(1) オンラインと実世界の二層性に焦点を置く、(2) マイクロワールドとしてメッセージ付きジレンマゲームを用いる。(3) 複数エージェント間での同期の生起とフェイクニュース拡散の関連を検討する。

2. 先行研究

人間の言語は現実世界における実際の距離とは無関係に、回数の制限もなく情報を拡散させる。これにより、個人が知覚する周辺世界の現実との乖離が生じる。冒頭で述べたインターネット上の問題は、こういった言語の性質が引き起こすオンライン(サイバー世界)とオフライン(物理世界)の二重性の問題と考えられている (Mercier, 2020)。

上記のような二重性を実験室にて再現する研究として、メッセージ付きジレンマゲームにおけるコミュニケーションの研究 (Inoue & Morita, 2021) が挙げられる。この研究は、集団内での記号の発生を実験的に捉える実験記号論の流れを汲む。実験記号論に基づく実験パラダイムにおいて、プレイヤーは、通常のコミュニケーション手段が制限された状態で、他のプレイヤーとの協調ゲームを繰り返す。協調を達成するためには、意思疎通の手段が必要となる。この必要により、実験記号論に基づく研究では、ゲームの進行に伴う新たな記号体系の出現が観察される。

実験記号論は、インターネットを含むあらゆる人間のコミュニケーション状況に適用可能な抽象化されたマイクロワールド環境である。言語、すなわち集団内の記号体系は、成員間の信頼関係によって支えられ、集団に固有であり、かつ変化を前提としている。Inoue & Morita (2021) は、そのようなコミュニケーションの変化にフォーカスし、通常の実験記号論における協

調ゲームに、裏切りや競争の要素を含めたメッセージ付きジレンマゲームを提案した。

メッセージ付きジレンマゲームは、環境内の資源を探索し、それを他のプレイヤーと同時に取得する、もしくは単独で取得することによって得点を得る課題である。他者との協調により、環境探索のコストが軽減され、より多くの報酬を取得することができる。しかし、コミュニケーションにかかるコスト次第で、単独で報酬を取得することのほうが効率がよい状況もありえる。また、他者との協調関係を維持したまま、単独で報酬を取得することができれば、より多くの報酬が得られる可能性もある。この可能性は、誤情報による他者の誘導、すなわちフェイクニュースの発信として実現できる。そして、誤情報による誘導は、単独で得る報酬が協調で得る報酬に比べて顕著に大きく、かつ利用される記号体系に、裏切りの意図を隠蔽する曖昧性が存在するとき生起すると考えられる。

上記のようにメッセージ付きジレンマゲームの構造を操作することで、構築される記号体系とプレイヤー間の人間関係に関する多様な状況をシミュレーションできる。本研究では、メッセージ付きジレンマゲームのこういった性質を利用し、ローカルエコチェーンの形成の背後にある認知機構を分析する。その際、コミュニケーションの形成に共通する一般的な認知機構として、事例ベースの学習理論 (IBLT, Gonzalez et al. 2003) を仮定する。つまり、過去のうまくいったコミュニケーション事例を繰り返すことで、送信する他者に送信するメッセージを選択し、解釈し、行動を決定するエージェントを構築する。

3. モデル

著者らは、コミュニケーションにおけるローカルエコチェーンは、エージェントの配置に局所性を持たせた試行が蓄積されることで出現すると考えている。IBLTに基づくエージェントは、隣接するエージェントと共通の経験を蓄積し、特定のエージェントのクラスター内でのみ共有される記号体系に至る。クラスター内で共通する記号体系は、クラスター外のエージェントを排斥する原因となる。こういった極化のプロセスを実現するために、IBLTを実装したPythonライブラリであるPyIBL^{*1} (Morrison & Gonzalez, 2024) をプラットフォームとした。このライブラリは、ACT-R (Anderson, 2007) などを用いた従来の認知モデルに比べ、高速な処理による大規模なIBLエージェントを用いたシミュレーションを可能とする。

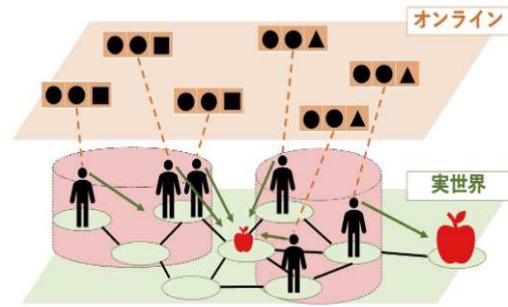


図 1: モデル全体の構想図

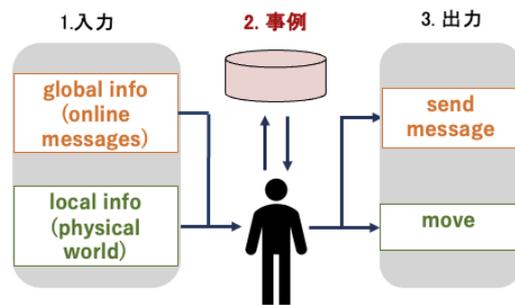


図 2: IBL エージェントの意思決定

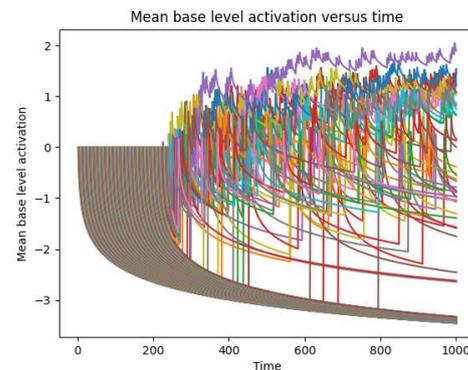


図 3: Inoue & Morita (2021) を模した課題において得られた各事例のベースレベル活性化値の推移

モデルの全体の構想を図 1 に示す。メッセージの交換が行われるオンライン、エージェントが実際に移動することが可能な実世界の 2 つの層が存在する。各エージェントはオンライン層の全てのメッセージを閲覧することが可能だが、実世界層の情報は自身の現在いる部屋とそれに隣接する部屋のもののみ観測可能である。図 2 で示すように、エージェントはグローバルな情報（オンライン層のメッセージ情報）とローカルな情報（実世界の部屋情報）を入力として受け取り、次に送信するメッセージと行動先を決める。

IBLT に基づいた意思決定を行うため、エージェントは、過去の事例を文脈に応じて連想することが可能である (図 2)。各事例はその思い出しやすさに影響

^{*1} <http://pyibl.ddmlab.com/>

する活性値を保持しており、行動に応じて報酬を与えることで活性値を操作する。図3で示すように、活性値は事例を連想することで上昇し、使用されない事例の活性値は時間の経過と共に減少する。これを繰り返すことで、エージェントは経験を通して獲得したメッセージと行動結果を学習し、先行研究と同様にコミュニケーションシステムを確立することが期待される。

提案するモデルにおいては、局所的に異なるコミュニケーションシステムを持つローカルエコーチェンバーが生成されること、エージェント間で誤情報の送信などにより得点の独占が生じる過程を観測する。

4. 予備的シミュレーション

4.1 目的

前節のモデルを用いたシミュレーションの前段階として、単純化された状況でのIBLエージェントの振る舞いを検討した。この予備的シミュレーションの設定では、環境内に資源を配置しなかった。エージェントのみが存在する環境で、エージェント同士が出会うことで報酬を獲得することとした。ただし、ゲームに参加するエージェントは3体以上とし、2体以上のエージェントの部屋の一致をスコア獲得の条件とした。これにより、特定エージェント間のクラスタが形成され、クラスタ外のエージェントとの間で獲得報酬に差が生じる可能性がある。

4.2 方法

グリッドサイズ3*3の環境にて、3体のエージェントがメッセージを送受信し、移動した。このステップの結果として、2体のエージェントの位置が一致した場合に、3体のエージェントの位置はランダムに配置され直した。各エージェントは、6種類の図形を3つ組み合わせた系列をメッセージとして送信した。メッセージの送信順序は、3体のエージェントの間でランダムであった。送信されたメッセージは、全エージェントに共有された。なお、この予備シミュレーションで、エージェントは周辺環境を知覚することはできず、自身の位置と自身のメッセージ決定以前に受信した他プレイヤーからのメッセージに基づいて送信メッセージを決定した。そして、3体のエージェントがメッセージを送信し終えた後、自己の初期位置と受信されたメッセージを条件として行動を決定した。

提案手法であるIBLTモデルでは、各ステップの行動決定の結果が記憶として保持され、続くステップにおいて利用された。本研究では、このIBLモデルをランダムに移動するモデルと比較した。それぞれのモデルは、100stepのシミュレーションを5回実行した

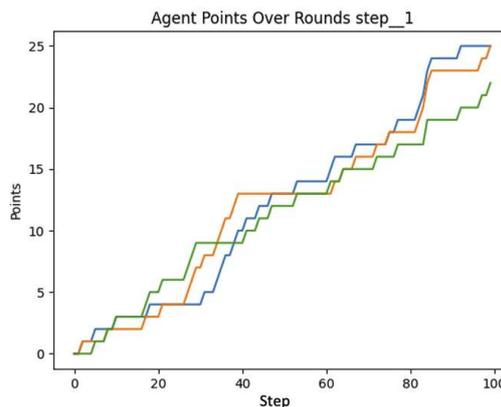


図4: チェンバー化が生じないシミュレーションの事例 (エージェント3体の得点の推移)

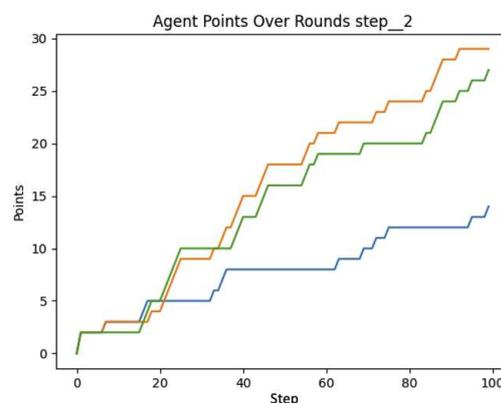


図5: チェンバー化が示唆されるシミュレーションの事例 (エージェント3体の得点の推移)

4.3 結果・考察

エージェント単体の平均獲得スコアはIBLモデル=24.27, ランダムモデル=20.00となった。これらの間には有意な差が見られた ($t(27) = 4.7, p < 0.01$)。この結果はIBLエージェントがランダムで意思決定を行うエージェントと比較して獲得スコアが高く、IBLTに基づいた学習が実際に行われていることを示す。ここからは、IBLエージェントの間に、何らかのコミュニケーションシステムが形成されたことが示唆される。

本研究の目的であるローカルエコーチェンバーの生起を検討するために、具体的なシミュレーション事例を検討した。エージェント3体でのシミュレーションにおけるスコアの推移を分析した結果、エージェント間のスコアに差が生じなかった事例と、差が生じた事例が確認された。図4はスコアに差が生じなかった事例、図5はスコアに差が生じた事例であり、どちらもIBLモデルの結果である。図4ではエージェント3体全員が近いスコアの推移を示しており、エージェント同士が偏りなく接触していることが推測できる。したがって、この事例においてはチェンバー化が進行して

いないことが示唆される。図5では2体のエージェントが近いスコアの推移を示している一方、1体のみ他2体とは異なる推移を示しており、エージェントの接触に偏りが見られる。これは、この事例においてはチェンバー化が進んでいる可能性を示唆する。

詳細なメッセージの分析は将来的な検討を待たねばならないものの、チェンバー化が示唆される事例では、チェンバー内のエージェントのみで共有される記号体系が形成された可能性がある。また、この記号体系は、クラスター外のエージェントにとって、誤情報として機能したと解釈することもできる。

5. まとめ・今後の課題

本研究ではオンラインコミュニケーションに対する信頼の再構築に向けて、メッセージ付きジレンマゲームをIBLTに基づいて拡張するモデルの構想を議論し、ローカルエコーチェンバー形成メカニズムとフェイクニュース拡散のシミュレーションを行うための枠組みを示した。

本研究で示した予備的シミュレーションは、提示された枠組みを実現するものではないものの、エージェント間の局所化の事例は示した。今後、本研究では省略されたエージェントの機能（環境情報の取得）の実装および環境を操作するシミュレーションを実施することで、ローカルエコーチェンバーの生起に関する検討を進めていく予定である。将来的には、シミュレーションの規模を拡大し、ネットワーク分析などの手法を用いたより詳細な分析が求められる。

6. 謝辞

本研究は、JST RISTEX JPMJRS23L3の支援を受けたものである。

文献

- Anderson, J. (2007). *How can the human mind occur in the physical universe?:* Oxford University Press.
- DDML (2024). *PyIBL 5.1.4 Manual*, URL: <http://pyibl.ddmlab.com/>, Accessed: 2024-04-26.
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). "Instance based learning in dynamic decision making." *Cognitive Science*, 27(4), 591–635.
- Inoue, N., & Morita, J. (2021). "A behavioral task for exploring dynamics of communication system in dilemma situations." *Artificial Life and Robotics*, 26(3), 329–337.
- Mercier, H. (2020). *Not born yesterday: The science of who we trust and what we believe:* Princeton University Press.
- Vosoughi, S., Roy, D., & Aral, S. (2018). "The spread of true and false news online." *Science*, 359(6380), 1146–1151.
- 森田純哉 (2023). ローカルエコーチェンバーをステアリングするトラスト調和メカニズムの認知的検討, URL: <https://www.jst.go.jp/ristex/digist/projects/r05-pi-b01.html>, Accessed: 2024-04-26.
- 総務省 (2019). インターネット上での情報流通の特徴と言われているもの, URL: <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/html/nd114210.html>, Accessed: 2024-07-19.