

# 同調するロボットが参加するスモールワールドネットワーク の意見伝播に関する社会シミュレーション

## Social Simulation of Opinion Dynamics in Small-World Networks Including Conforming Robots.

布施 陽太郎<sup>1</sup> \*

Yotaro Fuse<sup>1</sup>

<sup>1</sup> 富山県立大学

<sup>1</sup> Toyama Prefectural University

**Abstract:** ロボットに人らしさや社会性を実装するために、他者の意見に同調するロボットが研究されている。また、幾つかの実験的研究がロボットとのやり取りは人の意見を変容させることを報告している。そのようなやり取りが社会の各所で発生し人々の意見を変容させることが社会に与える影響についてほとんど検証されていない。スモールワールド性を持つ生成された人的ネットワークに新たなノードとしてロボットを追加することにより人口ロボット共存社会を表現し、そのネットワーク構造での意見伝播の特性をシミュレーションによって調査する。

### 1 はじめに

人とロボットの相互作用に関する研究は、人とロボットが共存する社会の実現を目的として現在進められている [1, 2]. 特に、人らしく相互作用するロボットはソーシャルロボットと呼ばれ、人にとって受容可能なソーシャルロボットの在り方が探求されている。ロボットが人々に受け入れられるためには、ロボットが社会規範に適応することが重要である。社会規範は集団構成員によって共有され、その集団内での行動志向の意思決定に情報を提供し、人々の健康と幸福に影響を与える可能性がある [3]. さらに、それらは人々が共に暮らすことを助け、協調性と社会秩序を促進する [3]. 人社会では、期待される行動は常に明示的に述べられるわけではなく、暗黙のうちに共有されることが多い。そして、それらの規範に適応する能力は、その人が集団の適切なメンバーであるかどうかを判断する基準となり得る。挨拶やジェスチャーなどの一部の社会規範は文化圏内で共有されているが、独自の集団規範は、複数のメンバーで構成される集団内で共有されている。先行研究では、暗黙のうちに形成された集団規範に従って適応し行動するエージェントやロボットの実証研究を実施しており、3人または4人の小規模な集団に焦点を当てている [4, 5].

CASA 理論 (Computers Are Social Actors) や Media Equation は人がコンピュータ、ロボット、その他のデ

ジタルデバイスを社会的な存在とみなしたり、メディアやそのコンテンツを実際の人や社会的な相互作用と同じように扱ったりする傾向があることについて指摘している [6, 7]. Thellman らのサーベイは、人々はロボットを計算機としてよりも擬人化し、ロボットよりも人として扱う傾向があることを指摘している [8]. また、自然言語処理の分野を劇的に変えつつある大規模言語モデルが、社会ロボットの対話能力を向上させる可能性が示唆されている [9]. したがって、人のようなロボットや社会性のあるロボットと人との共存がますます進む未来社会を考えると、そのようなロボットと人との間にも社会的影響が観察される可能性がある。人の行動や意見に変化をもたらす人同士の社会的影響は、他者の存在や他者との相互作用による部分がある [10]. したがって、人々が社会的影響を及ぼし合うように、人に擬人化されたロボットと人との相互作用においても、人とロボットの社会的影響が起り得る。実際にいくつかの先行研究が実証実験において社会的なロボットは人の行動や考えに影響を与えることを示している [11, 12, 13, 14, 15, 16, 17]. 加えて、ロボットが人に及ぼす社会的影響は、ロボットの外見、知能、社会的地位、人がロボットを受け入れる程度など、さまざまな要因が複雑に組み合わさったものとなる可能性がある。

加えて、AI 倫理の観点からロボットの振る舞いが人の意思決定や考えに影響を与えることを前提とし、ロボットが社会規範に従って意思決定することについての問題点が指摘されている [18]. 特に、ロボットが準拠

\*連絡先: 富山県立大学情報工学部データサイエンス学科  
〒939-0398 富山県射水市黒河 5180  
E-mail: fuse@pu-toyama.ac.jp

する社会規範が時代遅れになることやロボットによって共有されている既存の社会規範が変容することが定性的に指摘されている。そのような定性的な指摘に定量的データによる裏付けを加えることができれば、さらにその主張に関する議論が活性化する可能性がある。

一般的に、人社会はさまざまな規模の集団から構成されており、個人は複数の集団に属していることが多い。そのため、特定の集団への適応が人々の行動パターンに影響を与える。人々は主に、他者の行動や考え方が正しいと信じたり、他者に好かれたいという理由から同調するが、これにより、人々の行動や信念は変化しうる可能性がある [19]。さらに、その人が所属するさまざまな集団内での行動に影響を与える波及効果をもたらす可能性がある。社会規範を考慮するロボットの研究では、意思決定モデルに従ったロボットの行動や、小さい集団内で起こる現象のみが観察されている [4, 5]。しかし、多数の人とロボットが同時に参加する実験シナリオを設計することは時間や費用のコストの面から容易ではない。また、現在の人社会では、高度に社会化されたロボットは普及しておらず、実データを用いて分析することはできない。そのため、ロボットの同調が人社会に与える影響を実験的に検証したり、実世界からのフィードバックを取得して分析したりすることは現状困難である。

先行研究では意見伝播に関する社会シミュレーション手法を用いて、人とロボットの社会における個々のロボットが集団規範に適応しようとするのが社会全体に与える影響をシミュレートしている [20]。しかしながら、その社会シミュレーションでは、人とロボットの社会における社会構成員同士の繋がりがトラス状の格子ネットワークによって表現されている。トラス型格子ネットワークの構造ではすべてのノードが同一の接続パターンや条件を持つため、エージェントの意思決定がマクロに与える影響を分析しやすくする。しかしながら、現実の人々の社会ネットワークはスモールワールド性を持つことが示唆されている [22]。したがって、スモールワールド性をもつ人的ネットワークに追加的にロボットが参入することが、現実にも即したネットワークになると考えられる。

したがって、本研究では、人を模倣するエージェントで構成されるスモールネットワークにロボットが参加する場合の意見伝播の社会シミュレーションを実施する。スモールワールド性を持つネットワークの生成手法として Watts–Strogatz モデル [21] を採用する。この研究では周囲の社会構成員に同調する振舞いをするロボットがネットワークに参加することが、社会全体での合意形成にどのような影響を与えるかを定量的に分析する。具体的には、シミュレートされた人とロボットの社会における多数意見と規範の形成に、ロボットの人に対する社会的影響力の強さ、他人の意見

に対する人の受容閾値、ロボットの数が与える影響を調査する。

人とロボットのインタラクション分野において、マルチエージェントシミュレーションを用いてロボットの社会的影響力を定量的に探求した研究はほとんどない。また、倫理分野では社会ロボットに関する定性的な議論が進められているが、合意形成シミュレーションの結果は、ロボットや人工知能 (AI) の存在が社会全体にどのような影響を与えるかについて、定量的なデータに基づく議論の土台を提供できる可能性がある。このような定量的なシミュレーションは、他者への共感という人的な特性を組み込んだロボットや AI チャットボットを設計するための基盤構築に役立つ可能性がある。また、シミュレーション結果は将来のインタラクション実験や実地研究に役立つ予備データを提供できる可能性がある。

## 2 方法

### 2.1 シミュレーション概要

人を模した多数のエージェント (HA: Human Agent) とロボットの代わりとして集団規範アウェアな意思決定をする多数のエージェント (RA: Robot Agent) が混在する環境を設定し、意見合意形成に関するシミュレーションを実施する。0 から 1 の連続値をとる意見  $x \in [0.0, 1.0]$  を各エージェントが持ち、ステップごとにエージェントが他のエージェントとインタラクションする際にある条件に基づいて両エージェントがそれぞれの意見  $x$  を更新する。人エージェント HA は合意形成シミュレーションで知られる Hegselmann–Krause モデル [24] で用いられる意思決定方法を、ロボットエージェント RA は我々の先行研究の意思決定手法を採用する。以上のシミュレーションによって、人エージェント HA とロボットエージェント RA で構成される社会における意見のダイナミクスをマルチエージェントベースで定量的に分析することを目的とする。

本研究ではロボットから人への影響力に関するパラメータを変化させることによる社会での合意形成への影響を調査する。人同士はお互いに意見について影響しあうが、ロボットから人への影響力 (影響係数) はどの程度の大きさになるのかは、ロボットの外見や知性のデザインに依存すると仮定する。したがって、人から人への影響力を基準にして、ロボットから人への影響力は同程度かそれよりも大きいもしくは小さいと仮定する。これによって、ロボットがどの程度の人の意見への影響力を持つと、社会全体での意見収束や合意形成に影響を与えるのかを調べることに繋がる。また、合意形成シミュレーションで用いられる Hegselmann–Krause

モデル [24] で一般的に用いられる他者の意見の許容閾値を変動させることによる影響も調査する。加えて、人の社会ネットワークへ参入するロボットの数も変動させる。これにより、社会の中のロボットの数が合意形成にどのような影響を与えるのかを調査する。よって本研究では、以上の3要因でパラメータを変化させることによる合意形成への影響をシミュレートする。

## 2.2 モデル

本研究では、Watts–Strogatz モデル [21] によって生成されたスモールワールド性を持つ社会ネットワークに配置された人エージェントとロボットエージェントのダイナミクスを観察する。スモールワールド性とは、非常に多くの個体（ノード）が存在する大規模なネットワークにおいても、ほとんどのノード同士がごく短い経路（リンクの数）で繋がっているという性質を指す [23]。先行研究ではトラス形状のグリッドネットワークが採用されている。そのネットワークは各エージェントが均質なつながりを持ちネットワーク構造に依存しにくいエージェント間の影響関係の分析には有用であった。しかし、現実の人々の社会ネットワークはスモールワールド性を持つとされており、その性質を再現するネットワークを生成する Watts–Strogatz モデルを本研究では採用する。

加えて、ロボットや AI ボットはスモールワールド性を持つネットワークに組み込まれるというよりも、そのネットワークに新たにノードとして追加される存在であると仮定する。すなわち、ある人がロボットを手に入れ、そのロボットがその人とその周囲の人々とつながりを新たに形成することを想定する。自然発生的な社会ネットワークに新たにロボットを導入することが、現実の人口ロボットネットワークに即したシミュレーション用のネットワーク生成につながると考えられる。

図 1 とアルゴリズム 1 にスモールワールド性を持つ人エージェントの社会ネットワークにロボットエージェントを追加するための過程を示す。アルゴリズム 1 に従って生成された人社会ネットワークにロボットを追加することによって、人口ロボット社会ネットワークを生成する。このアルゴリズムによって、人社会ネットワーク  $N_H$  に  $n_R$  体のロボットエージェントとそのエージェントとのつながりが追加された人口ロボット社会ネットワーク  $N_{HR}$  が生成される。

## 2.3 エージェント

本実験でネットワークに配置されるエージェントは人エージェント (HA) とロボットエージェント (RA) である。HA は他のエージェントとのやり取りの中で、意

---

### Algorithm 1 人口ロボットネットワーク $N_{HR}$ 生成手順

---

- 1: 人エージェント  $H$  のみによって構成されるネットワーク  $N_H$  を生成 (図 1 左参照)
  - 2: **for**  $i = 1$  to  $n_R$  **do**
  - 3:    $N$  を構成するある人エージェント  $H_n$  をランダムに選択 (図 1 中央参照)
  - 4:    $H_n$  と  $H_n$  の近傍のエージェントで構成される集団  $G_n$  を導出
  - 5:   集団  $G_n$  の各エージェントにロボットエージェント  $R_n$  を接続 (図 1 右参照)
  - 6: **end for**
- 

見が大きく異なるエージェントの意見を基にして自身の意見を更新することはない。しかし、RA はつながりのある周囲のエージェントの多数派意見に同調することを目的とする。社会規範への適応や同調は人にとってはその人がその社会に適応することの手段であり、場合によっては敢えて社会規範に適応しないということがあり得る。一方で、ロボットは人に受容されるために、その場で共有されている社会規範に適応することを手段とする。したがって、人とロボットの間の他者への同調の質的差異が HA と RA にはある。

### 2.3.1 人エージェント

人エージェント (HA) の意見更新方法は Hegselmann–Krause モデルを基にしている。Hegselmann–Krause モデルは、個々人が持つ意見のダイナミクスや社会的影響に関するマルチエージェントベースのモデルの 1 種である [24]。本モデルは個人の間の相互作用がどのように社会全体の意見の分布やコンセンサス形成に影響するかを理解するために用いられる。Hegselmann–Krause モデルには意見  $x$  を持つエージェントが存在し、各エージェントの意見の初期化の後、ステップごとに各エージェントの周囲のエージェントの意見を参考に自身の意見を更新することが繰り返される。そして、システムが安定する（意見の変化がなくなる）までステップでの更新を繰り返す。

Hegselmann–Krause モデルベースのエージェントの意見更新の手順を以下に示す。あるエージェント  $i$  の意見  $x_i$  は実数値  $[0.0, 1.0]$  によって表現される。意見の初期化では各エージェントの意見に一様分布に従う乱数が代入される。その初期ステップから、ステップを経るごとに各エージェントは一斉に意見の値を更新する。あるエージェント  $i$  が意見を更新する方法は式 1 と式 2 に示される。式 1 はエージェント  $i$  の意見  $x_i$  が意見  $x_i$  と  $i$  の周囲のエージェントが持つ意見の集合である  $G(i)$  の要素の重みづけ平均によって算出される。式 2 はその重みづけのための条件を示す。ある人エー

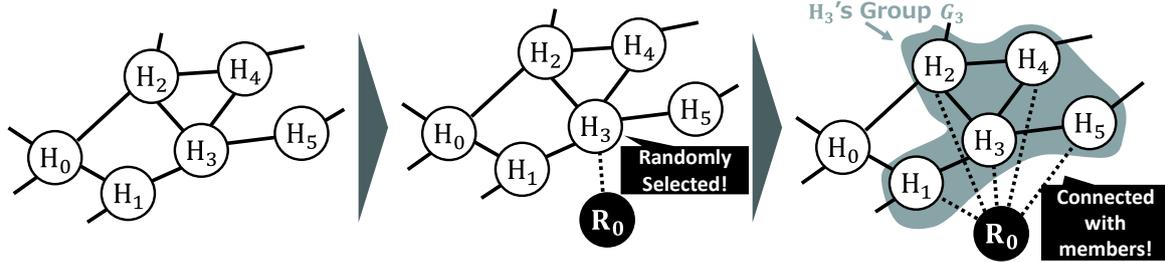


図 1: 形成されていた人エージェントのみで構成されるネットワーク (左) に 1 体のロボットエージェントを追加する流れ. ランダムに選ばれた人エージェント  $H_3$  がロボットエージェント  $R_0$  の所有者と見なされ (中央),  $H_3$  の近傍の人エージェントと  $R_0$  はつながりを生成する (右). このことは, 所有者の  $H_3$  とその周辺の人エージェントの集団メンバとして  $R_0$  が参加することを意味する.

エージェント  $j$  と  $i$  の意見の差が許容閾値  $\epsilon$  より大きければ, 重み  $w_j$  は 0 となり, 人エージェント  $i$  の意見更新に影響を与えない. これは人エージェント  $i$  からみて人エージェント  $j$  は意見がかけ離れており, 意見  $x_j$  を無視することを意味する. 許容閾値  $\epsilon$  以下である場合,  $i$  と  $j$  の意見は近く, お互いの意見が同意に近づくように意見  $x_i$  が調整されることを意味する.  $e_{R \rightarrow H}$  はロボットエージェントから人エージェントの意見への影響効率係数である. Hegselmann-Krause モデルにおいて人エージェント同士が影響を与える効率  $e_{H \rightarrow H}$  を 1.0 と定義する. ロボットから人への影響は必ずしも人同士の影響と同じ程度を持つとは限らないと仮定し, 制御パラメータとして  $e_{R \rightarrow H}$  を変化させて複数シナリオでのシミュレーションを実施する.

$$x_i = \frac{\sum_{j \in \mathcal{G}(i)} w_j x_j}{\sum_{j \in \mathcal{G}(i)} w_j}. \quad (1)$$

$$w_j = \begin{cases} 1, & \text{if } j \text{ is a HA and } |x_j - x_i| \leq \epsilon, \\ e_{r \rightarrow h}, & \text{if } j \text{ is a RA and } |x_j - x_i| \leq \epsilon, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

### 2.3.2 ロボットエージェント

先行研究 [4] で提案されている集団規範アウェアなロボットエージェント (RA) はステップ  $t$  において, 環境を観察し, その観察を基に内部状態を更新し, その内部状態に基づいて行動を出力する. そのような入出力をステップごとに実施し, 他者を観察しながら, エージェント自身の振舞いを集団に適応させ続けることを試みる. 本研究では, 観察対象は近傍の他者の意見  $x$  の集合  $X_{\text{nbd}}$  であり, 行動は自らの意見  $x_{\text{self}}$  である. 本研究で実施されるシミュレーションにおいて, 集団規範アウェアなエージェントの意思決定の方法は先行研究で提案された意思決定手法を基にして実装される.

合意形成シミュレーションにおいて, 意見  $x$  は  $[0.0, 1.0]$  の実数値を取る. RA は選択可能な意見の集合  $X_{\text{self}} = \{k\delta \mid k \in \mathbb{N}, 0 \leq k\delta \leq 1\}$  から 1 つの意見  $x_{\text{self}}$  を採用して出力する. 内部状態としてエージェントは価値関数  $V$  を持つ. エージェントは特定集団で暗黙的に共有されている規範に則った適応的行動を推定するために価値関数を用いる. 価値関数は, エージェントが実施可能な振舞いや準拠可能な行動実行の基準の集合に, 特定の集団内でそれに基づいて実際に意思決定することの価値の集合を割り当てる.

本意思決定手法は集団内の規範が存在することを前提として, 選択可能な意見集合  $X_{\text{self}}$  に対して, 意見  $x$  を表明する価値の集合を割り当てる. そして, その関数の中で最も高い価値を割り当てられた意見  $x^* = \arg \max V(x)$  がエージェントの意見として採用される. 価値関数は式 3 と式 4 によって更新される.

$$V(x) \leftarrow (1 - \alpha)V(x) + \alpha(R(x) + \gamma \max V(x)) \quad (3)$$

$$R(x) = \sum_{x' \in X_{\text{nbd}}} \left\{ \exp \left( -\frac{(x - x')^2}{\text{kurtosis}} \right) \right\} \quad (4)$$

すなわち, 各 RA は  $[0.0, 1.0]$  の範囲をとる意見  $x$  を複数個の実数値に等間隔に分割して選択可能な意見  $x' \in A$  として保持し, 各ステップにおいて価値関数に基づいて  $A$  の中から 1 つ選択しその RA 自身の意見  $x$  を採用する. 価値関数の更新に伴い RA が持つ意見が変遷していくことになり, このことは集団規範アウェアにロボットの意見を調整することを意味する.

## 2.4 パラメータ

表 1 には固定されたパラメータが示されている. 各エピソードにおける初期意見の分布は seed 値に基づいて統一されており, 制御パラメータの変化によって発生する変化のみを原因としたマクロへの影響を観察する.

表2に基づき、制御パラメータとして  $n_R$ ,  $\epsilon$ ,  $e_{R \rightarrow H}$  を変化させた混合モデルを用いて、複数のシナリオをシミュレーションする。1度のシミュレーションの実施において、 $n_R$ ,  $\epsilon$ ,  $e_{R \rightarrow H}$ , seed 値に個別の値が代入される。また、seed 値は0から49の整数値を取ることから、 $\epsilon$ ,  $e_{R \rightarrow H}$  によって定められる人工社会において、各エージェントが持つ初期意見とネットワーク構造が50パターン生成されることを意味する。すなわち、seed 値を50個用いてシミュレートすることで、初期意見とネットワーク構造に依存しない傾向を観察することを目論んだ。

意見閾値の大小は人エージェント社会全体での合意形成のしやすさに影響する。 $\epsilon$ を0.5もしくは0.25に設定することで、HAが全体的な合意形成を起しやすいつい人工社会と起しにくい社会でのシミュレーションを実施する。ロボットエージェントの数  $n_R$  は人口ロボット社会ネットワークに参入するロボットの数を意味する。特に、本研究においてロボットは人に受容されるために他者への同調を目的とする。この集団規範への同調というロボットの数が社会全体で増えていくことが社会での合意形成にどのような影響を及ぼすのかを評価する。加えて、パラメータ  $e_{R \rightarrow H}$  を制御することで、人の他者に対する受容の度合いやロボットから人への影響力の違いがマクロであるエージェント社会全体に与える影響を評価する。

表 1: 設定パラメータ.

Model Parameter	
Opinion $x$	[0.0, 1.0]
Number of Episodes (seed value)	50
Max Steps MAX	500
Watts–Strogatz Model Parameter	
Number of Humans $n$	100
Number of nearest neighbors $k$	4
Rewiring Probability $p$	0.1
Robot Agent Parameter	
Increment of Opinion $\delta$	0.01
number of set of actions $ X_{\text{self}} $	101
Learning rate $\alpha$	0.9
Discount factor $\gamma$	0.1
Gaussian parameter kurtosis	0.01

### 3 結果

図2, 3は各  $\epsilon$  条件と影響係数  $e_{R \rightarrow H}$ , ロボットエージェント (RA) の数  $n_R$  の組み合わせごとの最終ステッ

表 2: 制御パラメータ.

Number of Robots $n_R$	0, 25, 50, 100
Tolerance Threshold $\epsilon$	0.25, 0.5
Effect $e_{R \rightarrow H}$	0.01, 1.0, 10.0
Seed Value	0 to 49

プでの多数派の意見の変動度合いを示す。すなわち、人社会 (HS) と人口ロボット社会 (HRS) の間の中央値の差の絶対値  $|x_{\text{HS}}^M - x_{\text{HRS}}^M|$  を示している。

ここでは同じ seed 値とパラメータの組み合わせの基で HS と HRS での最終ステップでの中央値を多数派意見  $x_{\text{HS}}^M$ ,  $x_{\text{HRS}}^M$  とみなして比較する。人エージェント (HA) で構成される社会ネットワークに RA を追加する形でシミュレーションを実施した。そのため、RA が  $n_R$  体追加された場合と0体追加された場合で比較することによって、RA が参加することによって多数派意見がどのように変動したのかを算出する。すなわち、HA のみで構成される社会での最終的な多数派意見  $x_{\text{HS}}^M$  と比べて、多数派に同調する RA が混在する社会での最終的な多数派意見  $x_{\text{HRS}}^M$  はどのように変化するかを観察する。このことは、同調 RA がミクロである集団へ社会的に適応することが、マクロである人口ロボット社会全体に影響を与えることに関する指標となる。

縦軸は HRS での最終的な意見の中央値から人社会での最終的な意見の中央値を引き算した値の絶対値である。すなわち、縦軸の値が0より大きい場合、最終的に意見が合意ないし収束に至った時に多数派意見として採用された値において HS と HRS で変容があることを示す。横軸は影響係数  $e_{R \rightarrow H}$  とロボットの数  $n_R$  である。

ただし、 $\epsilon = 0.25$  条件のいくつかの社会では合意される意見が2つに分化し、どちらの意見も社会全体で過半数のエージェントからの支持を得られなかった可能性がある。そのような社会における合意された意見を中央値によって算出することは不適切であり、意見が分化した社会についてのより詳細な分析が必要である。

### 4 考察

図2, 3より、ロボットから人への影響に関する係数  $e_{R \rightarrow H}$  の値が大きくなるほど、 $|x_{\text{HS}}^M - x_{\text{HRS}}^M|$  も大きくなる傾向が観察されている。ただし、その傾向は  $e_{R \rightarrow H}$  条件の値の大きさにより異なる。人々が他者との意見の違いに寛容な場合 ( $e_{R \rightarrow H} = 0.5$ ), 社会全体での多数派の意見の変動は比較的大きくない、影響係数  $e_{R \rightarrow H} = 10$  という極端にロボットから人への影響が大きい場合であっても多数派意見の変動量は0.1程度である。

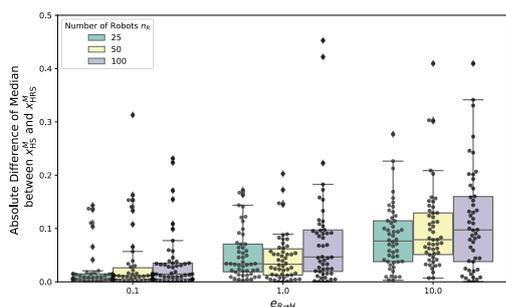


図 2:  $\epsilon = 0.25$  条件における人社会と人口ロボット社会での意見の中央値の差の絶対値。

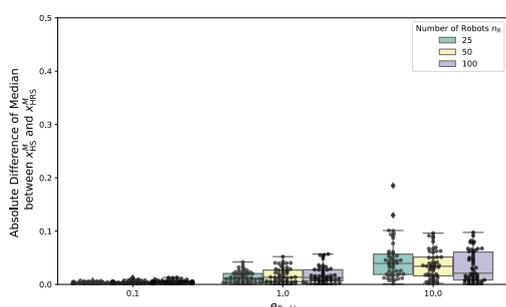


図 3:  $\epsilon = 0.50$  条件における人社会と人口ロボット社会での意見の中央値の差の絶対値。

一方、人々が他者との意見の違いに比較的不寛容な場合 ( $e_{R \rightarrow H} = 0.25$ )、であるロボットの影響力が人と同等の場合 (影響係数  $e_{R \rightarrow H} = 1$ ) においても、最大 0.2 の変容を起こす可能性が示唆されている。加えて、極端にロボットの影響係数が大きい場合 ( $e_{R \rightarrow H} = 10$ )、1 クォーター以上の変動を起こし得ることが示されている。ただし、これらの大きな変動は、2 つの意見の合意に分極化した社会で起きた可能性もあるので、詳細な分析が必要である。

興味深いことに、ロボットの数  $n_R$  の変化は影響係数条件内では、 $|x_{HS}^M - x_{HRS}^M|$  の変動に影響を与えないことが明らかになった。多数派に同調するロボットの数が社会ネットワークにおいて増加しても、人々の合意形成の在り方に影響を与えにくいことが示唆されている。ロボットの数  $n_R < 25$  かつ影響係数が小さくない場合 ( $e_{R \rightarrow H} \geq 1$ ) の条件においても、多数派意見の変動が大きくなり始めるロボットの数に関する条件を探索する必要がある。

以上のことは、人々が他者との意見の違いに不寛容かつロボットから人への影響力が大きい条件下では、人ネットワークに多数派へ同調する振舞いをするロボットが参加することがネットワーク内で多数派となる意見を大きく変える可能性があることを示している。ま

た、ロボットの数という条件はそのような多数派意見の変容に寄与しないことが明らかになった。すなわち、人が他者との意見の違いに不寛容であり、そのような他者の影響を受けないとみなされるトピック (例えば、政治や宗教、文化的振舞い) に関して、人がロボットの意見を参照しやすい場合、社会全体での意見の変容が大きく起こる可能性があることが示唆されている。

## 5 おわりに

本研究では集団規範アウェアな意思決定をするロボットエージェントが人社会に参入してインタラクションすることによる影響をシミュレーションによって評価した。特に、スモールワールド性を持つ人社会ネットワークにロボットがノードとして加わった人口ロボット社会ネットワークを生成し、その社会への影響がロボットの数やロボットの持つ社会的影響力、他者に対する意見許容値を制御することで探求した。社会を構成する人がロボットなどのエージェントから受ける影響を定量的にシミュレーションすることは、他者への同調という人が持つ特性を含めたエージェントデザインの土台構築に貢献することが期待できる。加えて、このようなシミュレーション研究の結果が実際のロボットや人を用いた大規模な実証研究をする際の指針となり得る。シミュレーション結果から、人が他者との意見の違いに不寛容であり、そのような他者の影響を受けないとみなされるトピック (例えば、政治や宗教、文化的振舞い) に関して、人がロボットの意見を参照しやすい場合、社会全体での意見の変容が大きく起こる可能性があることが示唆されている。将来的には、他のパラメータの制御、ネットワーク構造の変更による人口ロボット社会での合意形成を探索する。

## 参考文献

- [1] Leite, I. et al.: Social robots for long-term interaction: A survey, *International Journal of Social Robotics*, Vol. 5. No. 2, pp. 291–308 (2013)
- [2] Sheridan, T. B.: A review of recent research in social robotics, *Current Opinion in Psychology*, Vol. 36, pp. 7–12 (2020).
- [3] Legros, S. and Cislighi, B.: Mapping the social-norms literature: An overview of reviews, *Perspectives on Psychological Science*, Vol. 15, No. 1, pp. 62–80 (2020).
- [4] Fuse, Y. and Tokumaru, M.: Social influence of group norms developed by human-robot groups, *IEEE Access*, Vol. 8, pp. 56081–56091 (2020).

- [5] Fuse, Y., Ashida, B., Ayedoun, E., and Tokumaru, M.: Unleashing fairness: How a group norm-aware agent shakes up the ultimatum game, *IEEE Access*, Vol. 11, pp. 36727–36740 (2023).
- [6] Nass, C., Steuer, J., and Tauber, E. R.: Computers are social actors, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 72–78 (1994).
- [7] Reeves, B. and Nass, C.: The media equation: how people treat computers, television, and new media like real people and places, Cambridge University Press, USA (1996).
- [8] Thellman, S., de Graaf, M., and Ziemke, T.: Mental state attribution to robots: A systematic review of conceptions, methods, and findings, *ACM Transactions on Human-Robot Interaction (THRI)*, Vol. 11, No. 4, Article 41, pp. 1–51 (2022).
- [9] Zhang, C., Chen, J., Li, J., Peng, Y., and Mao, Z.: Large language models for human–robot interaction: A review, *Biomimetic Intelligence and Robotics*, Vol. 3, No. 4, Article 100131 (2023).
- [10] Spears, R.: Social influence and group identity, *Annual Review of Psychology*, Vol. 72, pp. 367–390 (2021).
- [11] Powers, A. and Kiesler, S.: The advisor robot: Tracing people’s mental model from a robot’s physical attributes, *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, pp. 218–225 (2006).
- [12] Ham, J. and Midden, C. J. H.: A persuasive robot to stimulate energy conservation: The influence of positive and negative social feedback and task similarity on energy-consumption behavior, *International Journal of Social Robotics*, Vol. 6, No. 2, pp. 163–171 (2014).
- [13] Chidambaram, V., Chiang, Y.-H., and Mutlu, B.: Designing persuasive robots: How robots might persuade people using vocal and nonverbal cues, *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, pp. 293–300 (2012).
- [14] Salomons, N., van der Linden, M., Strohkorb Sebo, S., and Scassellati, B.: Humans conform to robots: Disambiguating trust, truth, and conformity, *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 187–195 (2018).
- [15] Salomons, N. and Scassellati, B.: Trust and conformity when interacting with a group of robots, *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 315–316 (2018).
- [16] Qin, X., Chen, C., Yam, K. C., Cao, L., Li, W., Guan, J., Zhao, P., Dong, X., and Lin, Y.: Adults still can’t resist: A social robot can induce normative conformity, *Computers in Human Behavior*, Vol. 127, Article 107041 (2022).
- [17] Driver, T. B., Simecek, M., Clarke, C., You, Y., and Fraune, M. R.: Who is a better salesperson? People conform more to a human than robot facilitator in the wild, *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 427–431 (2024).
- [18] Coggins, T. N. and Steinert, S.: The seven troubles with norm-compliant robots, *Ethics and Information Technology*, Vol. 25, No. 2, Article 29 (2023).
- [19] Sunstein, C. R.: Conformity, NYU Press (2019).
- [20] 布施陽太郎: 集団規範アウェアなエージェントが人と共存する社会に関する合意形成シミュレーション, *HAI シンポジウム 2024*, P-11, 2024.
- [21] Watts, D. J. and Strogatz, S. H.: Collective dynamics of ‘small-world’ networks, *Nature*, Vol. 393, No. 6684, pp. 440–442 (1998).
- [22] Milgram, S.: The Small World Problem. *Psychology Today*, Vol. 2, pp. 60–67, (1967).
- [23] Watts, D. J., *Six Degrees: The Science Of A Connected Age*, W. W. Norton & Company (2004). (ダンカン・ワッツ, 辻竜平 (訳), 友知政樹 (訳): スモールワールド・ネットワーク [増補改訂版]: 世界をつなぐ「6次」の科学, 筑摩書房, (2016).)
- [24] Hegselmann, R. and Krause, U.: Opinion dynamics and bounded confidence: Models, analysis and simulation, *Journal of Artificial Societies and Social Simulation*, Vol. 5, No. 3 (2002).