

ソシオン理論に基づきモデル化したエージェントと 人との関係性のシミュレーション

The simulation of relations with people and a communicational agent
based on Socion theory

斉藤緑¹ 大隅俊宏¹ 大澤博隆² 村川賀彦³ 今井倫太²

Midori Saito¹, Toshihiro Osumi¹, Hirotaka Osawa², Yosihiko Murakawa³, and Michita Imai²

¹ 慶應義塾大学大学院理工学研究科

¹Graduate School of Science and Technology, Keio University

² 慶應義塾大学理工学部

² Faculty of Science and Technology, Keio University

³ 富士通研究所

³ FUJITSU LIMITED

Abstract: Recently, some agents like Roomba and Siri appear in human society. We have many more occasions to communicate between agents and people. Therefore, it is important to verify that human relations model can be adapted to agents whether or not. In this study we modeled agents and people based on Socion theory and simulated them to search changes in their relationships.

1 はじめに

近年 Roomba や Siri を始めとする様々なエージェントが人間社会に登場しており、将来はロボット、CG キャラクタ、AI キャラクタといった様々なエージェントと触れる機会が多くなると思われる。そのためにはエージェントも人間関係に適応できることが大切である。

従来研究では対話型ヒューマノイドロボットによる小学校のクラスでの友達関係の推定[1]や、エージェントが“好き”もしくは“嫌い”の1つの値に基づいて行動を取るシミュレータ[2]が作成されている。また、ロボットの社会性が人とロボットの関係や人間関係に与える影響をバランス理論を用いて調べた研究[3]がある。しかしながら大人数においては人間関係の把握に限界があり、実証実験にはコストがかかる。また、エージェントにも人間関係を表すためのモデルが適用できるかどうかを検証する必要がある。

そこで本研究ではソシオン理論[4]におけるソシオグラムによってエージェントをモデル化し、ソシオン理論を適用した人間との関係性を調べるためにシミュレーションを行った。

2 ソシオン理論

2.1 自分の定義

ソシオン理論では自己を3つの要素で定義し、視覚的に表現した図形をソシオグラムと呼ぶ。図1のように相手を思う自分(赤い矢印)、相手から思われる自分(青い矢印)、自分を思う自分(緑色の矢印)を使い相手との関係を表す。複数の自己を持つことで、より柔軟に社会的な関係性を表現できる。

2.2 自分と他者との関係

ソシオグラムでは、好き嫌いを示すために自分と相手とを結びつける結合強度として荷重と呼ばれる社会関係の基本単位を使用する。

図2において、白い荷重円は“好き”であることを、黒い荷重円は“嫌い”であることを示す。よってAさんはBさんが好きだが、BさんはAさんが嫌いということが表現できる。CさんとAさん、CさんとBさんについても同様である。

2.3 客観的・主観的な人間関係

ソシオン理論では人間社会と同じように、現実の

人間関係と頭の中で独自に想像した人間関係を表すことができ、図3のようにそれぞれCモード、Pモードと呼ぶ。CモードとPモードは必ずしも一致するわけではないため、人間関係の誤解やすれ違い、思い込みを表すことが可能である。図3をソシオグラムで表すと図4のようになる。青い荷重円で3者間に客観的な関係、つまりCモードがあることを、赤い荷重円で3者の頭の中に主観的な3人の関係、つまりPモードがあることを表している。これにより、客観的な関係、主観的な関係が互いにどのように影響するかがわかる。

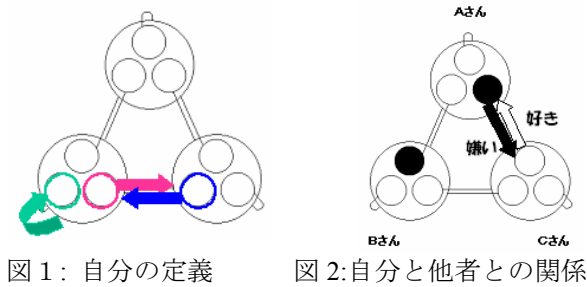


図1: 自分の定義 図2: 自分と他者との関係

その情報に影響を受けて他者を好きになったり嫌いになりながら人間関係を構築していく。このモデル[5]を使用し、人への変化する荷重や話題対象への荷重の情報、自身の荷重を更新する。

ここで、Cモードにおける人 a_i の他者 a_j に対する好感度は W_{ij} として表し、他者 a_k のPモードにおける人 a_i の他者 a_j に対する好感度は W_{ij}^k として表す。また、人 a_i のCモードの好感度と人 a_i のPモードの a_i が持つ好感度は同じなので $W_{ij} = W_{ij}^i$ である。

4.1 コミュニケーション相手の選択方法

今回、人は仲の良い相手、あるいは関心の高い相手とよく話し、気が合わない相手とはあまり話さないと考える。[5]のモデルでは人同士の仲の良さを表すパラメータを親密度とする。ある人 a_i と他者 a_j の新密度 f_{ij} は自分からの好感度と相手からの好感度の合計値に依存した(1)式で表す。コミュニケーション相手に選択する確率は(2)式で表す。

$$f_{ij} = \begin{cases} W_{ij}^i + W_{ji}^i & (W_{ij}^i + W_{ji}^i > 0) \\ 0 & (\text{otherwise}) \end{cases} \quad (1) \quad (j \neq i)$$

$$p_{ij} = \frac{f_{ij}}{\sum_{k=1}^n f_{ik}} \quad (k \neq i) \quad (2)$$

もし a_i が a_j を好きだとしても a_j が a_i を嫌いであった場合、親密度 f_{ij} は小さくなり、 a_j をコミュニケーション相手に選択する確率は低くなる。

3.2 話題とする人の選択方法

[5]のモデルでは、人は他者とのコミュニケーションにおいて、自分もしくはコミュニケーション相手にとって関心の高い対象を話題とすることが多いと考えている。今回は話題の対象は人間とし、話題提供をするのはコミュニケーション相手とする。コミュニケーション相手の各人に対する関心の高さに依存し話題となる人を決定している。

ある人 a_k が話題対象となる確率は(3)式で表す。

$$t_k^{ij} = \frac{|W_{jk}^i|}{\sum_{l=1}^n |W_{jl}^i|} \quad (l \neq i, l \neq j) \quad (3)$$

3.3 Pモードの更新

ソシオン理論のPモードに則り [5]のモデルでは、人は他者とのコミュニケーションを通じて自分のP

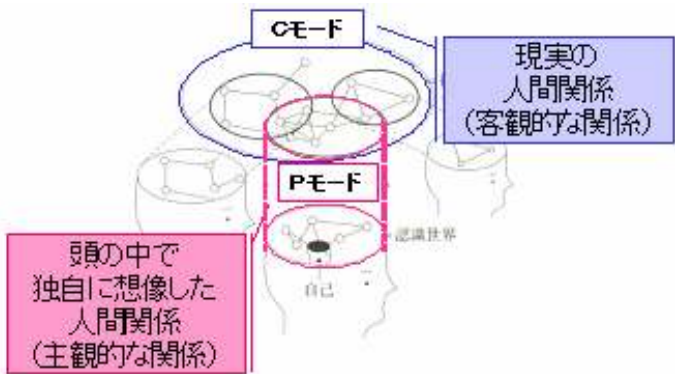


図3: 客観的・主観的な人間関係

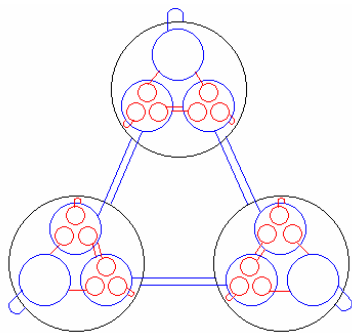


図4: CモードとPモードによる人間関係の同時表記

3 コミュニケーション方法

3.1 コミュニケーションモデル

人が他者と共にエージェントとコミュニケーションをする場面を考える。この時、コミュニケーションをしながら誰が誰を好きか嫌いかという情報を得、

モード内の“コミュニケーション相手の自分に対する好感度”と“コミュニケーション相手の話題対象の人に対する好感度”を更新する。また、コミュニケーション相手となった人もPモード内の“コミュニケーションしてきた人の自分に対する好感度”を更新する。CモードとPモードの不一致を表現するため、[5]のモデルではPモードの更新時に好感度の推定の誤差をゆらぎによって表現する。

ある人 a_i が他者 a_j をコミュニケーション相手として選び、 a_k を話題対象に選んだ場合、 W_{ji}^i と W_{jk}^i について(4)式と(5)式で人 a_i のPモードを、 a_i のコミュニケーション相手である a_j の W_{ij}^j を(6)式で更新する。

$$W_{ji}^i = W_{ji} + \alpha_1 \delta \quad (4)$$

$$W_{jk}^i = W_{jk} + \alpha_2 \delta \quad (5)$$

$$W_{ij}^j = W_{ij} + \alpha_1 \delta \quad (6)$$

δ は[-1,1]の一様分布の乱数であり、 α_1 はコミュニケーション相手から自分への好感度、 α_2 は自分以外への好感度を推定したときのゆらぎの大きさを表している。

3.4 コミュニケーション相手への好感度

更新

[5]のモデルでは、人はマンツーマンでのコミュニケーションにおいては互いに相手の好感度の影響を受けて自分の好感度を変化させる傾向が強く、同じ物を好きであったり、同じ物を嫌いであるほど互いの好感度が高まる傾向にあると考えている。そこで互いに相手の好感度の影響を受けて更新する部分と話題対象への好感度の差異の影響を受けて更新する部分の二段階によって互いの好感度を更新する。[5]のモデルでは、互いの関心の高さに依存した好感度を更新する確率を設定する。

ある人 a_i とコミュニケーション相手 a_j のコミュニケーションにおいて、 a_i と a_j が好感度を更新する確率はそれぞれ(7)式、(8)式で表す。自分の関心と比較してコミュニケーション相手の関心が高いほど好感度を更新する可能性が高くなる。

$$P_j^i = \frac{(W_{ij}^i)^2}{(W_{ij}^i)^2 + (W_{ji}^i)^2} \quad (7)$$

$$P_i^j = \frac{(W_{ji}^j)^2}{(W_{ji}^j)^2 + (W_{ij}^j)^2} \quad (8)$$

また、好感度を更新する場合の変化量は(9)式、(10)式で決定する。

$$\Delta W_{ij}^i = \omega_1 (W_{ji}^i - W_{ij}^i) \quad (9)$$

$$\Delta W_{ji}^j = \omega_1 (W_{ij}^j - W_{ji}^j) \quad (10)$$

また、話題対象への好感度の差異の影響を受けて更新する部分について人 a_i の好感度の変化量は(11)式で示す。

$$\Delta W_{ij}^i = \omega_2 (\chi - |W_{ik}^i - W_{jk}^i|) \quad (11)$$

χ は好感度の差異の閾値である。話題対象への好感度の差異が χ を下回っていればコミュニケーション相手への好感度が増加し、上回っていれば好感度が減少する。

3.5 話題対象への好感度更新

話題対象への好感度は、図5のようなバランス理論 (POX モデル) に基づいて更新される。実際の間人間関係において、自分Pと仲の良い友人Oがいた場合、図12(a)のようにOが好いている人Xとは仲良くなりやすく、(b)のようにOを嫌いな場合はOが嫌いな人Xとは仲は良くなりやすい傾向がある。一方、(c)のようにXを嫌っているOとは仲良くなりやすく、(d)のようにXを好いている人Oとは仲が悪くなりやすい傾向がある。人間は自分が好いている人が持つ好感度には大きく影響され、自分が嫌っている人が持つ好感度には影響されにくい傾向があると考えられる。そのため、[5]のモデルではコミュニケーション相手への好感度が正の場合のみ、話題対象への好感度をコミュニケーション相手に近づけるように更新する。

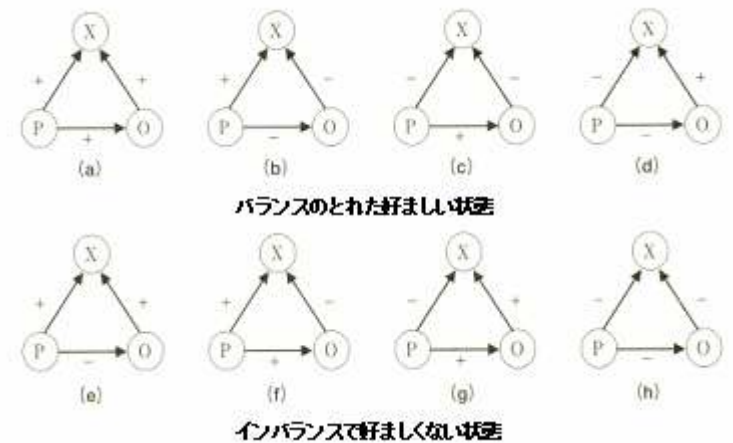


図5: バランス理論 (POX モデル)

ある人 a_i とコミュニケーション相手 a_j の話題対象 a_k において、話題対象への好感度の変化量は(12)式で表される。

$$\Delta W_{ik}^i = \begin{cases} \omega_3(W_{jk}^i - W_{ik}^i) & (W_{ij}^i \geq 0) \\ 0 & (\text{otherwise}) \end{cases} \quad (12)$$

3.6 地位向上方略

人間関係において人は他人よりも高く評価されたいという願望があり、[5]のモデルでは集団との同調によって自分の立場、評価が有利になるように立ち振る舞うことが他人の立場に影響を与えるという特徴があることを考えている。同調により自分の地位を向上させるために、人は自分自身の好き嫌いとは別に自分の行動によって自分の評価が上がるのか下がるのかを、自分のPモード内でシミュレーションし結果を予測した上で、自分が周囲から嫌われ者にならないように行動する傾向があると考えられる。バランス理論に基づいて周囲と同調することで自分の地位を保つ戦略である。そのため[5]のモデルでは、ある人 a_i の地位の高さ W_{ii} は他の人全員からの好感度の合計値に依存し、(13)式で計算する[6]。人はコミュニケーションによってPモードと好感度を更新した後、それぞれの人に対する好感度を上げた場合と下げた場合の自分の地位の変化量をシミュレーションする。ある人 a_i が他者 a_j の好感度を ΔW_{ij}^i 変化させた場合の人 a_j の地位の変化量 ΔW_{ii} は(14)式を用いる。

$$W_{ii} = \eta \sum_{j=1}^n W_{ji}^i = \eta(W_{1i}^i + W_{2i}^i + \dots + W_{ni}^i) \quad (13)$$

($j \neq i$)

$$W_{ii} = \eta \sum_{k=0}^n \omega_2 = (|W_{ij}^i - W_{kj}^i| - |W_{ij}^i - W_{kj}^i + \Delta W_{ij}^i|) \quad (14)$$

($j \neq i, k \neq i, k \neq j$)

(14)式は、各人が(11)式に基づいて好感度を更新しているため、自分と他者が持つ好感度の差異が小さいほど他者からの好感度が高くなることを表している。人はシミュレーションの結果、好感度を σ 上げた場合と σ 下げた場合それぞれにおいて自分の地位の向上が最も大きかった人を一人ずつ決定し、シミュレーション結果を反映している。

3.7 友人リンクと排斥リンク

人間関係では友人関係を構築したり、一方でいじめも深刻な問題である。そこで本研究では[7]のモデルを用い、友人関係になる場合といじめを受ける場合を考える。

・ある人 a_i の a_j に対する好感度 W_{ij} が友人閾値より大きいとき、 a_i から a_j に対し友人リンクを張る

・ある人 a_i の a_j に対する好感度 W_{ij} が排斥閾値より小さいとき、 a_i から a_j に対し排斥リンクを張る

これらによって構築された人間関係において、 a_x に対し以下の条件のどちらかもしくは両方が発生した場合、いじめが発生すると定義する。

- ・集団の半数以上の人から a_x に対する排斥リンクが張られている：集団全体によるいじめ
- ・互いに友人リンクを張っている3人以上全員から a_x に対する排斥リンクが張られている：グループによるいじめ

4 エージェントのモデル化

4.1 ソシオグラムのバリエーション

図6に示すように構造がないソシオンから構造を持つソシオンまで、ソシオグラムには様々なバリエーションが存在する。そのバリエーションによってソシオグラムで表現可能なエージェントを模索するために、定義できるソシオグラムのパターンの抽出を行った。それぞれ図7, 8, 9を参照し、定義できるソシオグラムの数は、

- ・他者へのCモードがない場合：2通り
 - ・他者へのCモードがある場合： $2^6 \cdot 2$ 通り
 - ・他者へのCモードが1つずつの場合： $2^3 \cdot 4$ 通り
- よって合計162通りとした。

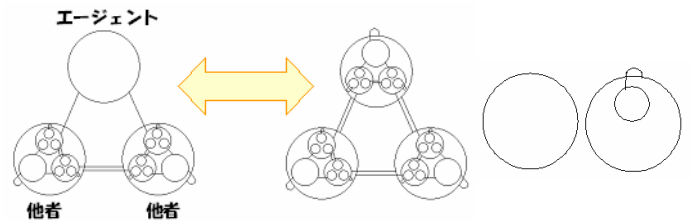


図6: バリエーションの範囲

図7: 2通り

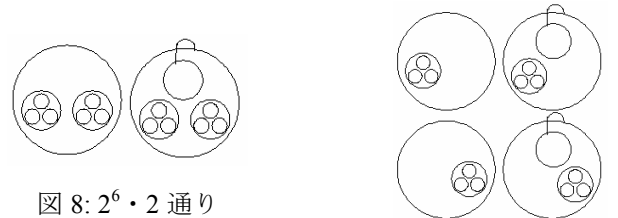


図8: $2^6 \cdot 2$ 通り

図9: $2^3 \cdot 4$ 通り

4.2 エージェントタイプ分類

抽出した162通りのソシオグラムのパターンの中からエージェントのタイプ进行分类する。本稿では例として3種類のタイプを選び、前節の式を使用してソシオグラムを表した。

4.2.1 エージェントタイプ 1

図 10 のように、ぬいぐるみや機械などは人間に働きかけることも、自身がどう思われるかも考えないため C モード, P モードの内部モデルをもっていないソシオグラムで表すことができる. (7), (8)式において $W_{ij}^i = 0, W_{ji}^i = 0$ としたため, (1)式によるコミュニケーション相手に選択されなくなってしまうため, 他者 a_j の C モード W_{ji}^j の値を $[-0.2, 0.2]$ の一様分布の乱数で更新させるようにした.

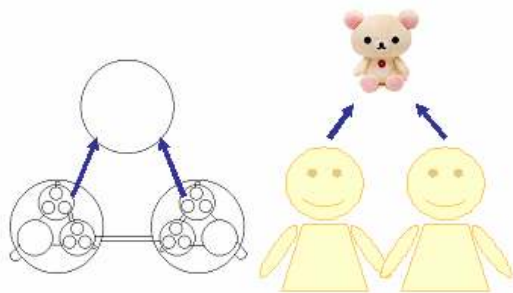


図 10: エージェントタイプ 1

4.2.2 エージェントタイプ 2

ヒューマノイドロボットと人とのインタラクションでは, 人の友達関係を推定するために ID タグを使って認識している. この場合はタイプ 2 として図 11 のように, 認識した人同士の互いの好き嫌いを, エージェントのソシオグラムの P モードによって表すことができる. (5)式を友達推定の式とし, (4), (7), (9)式については W_{ji}^i を更新せずに使用した.

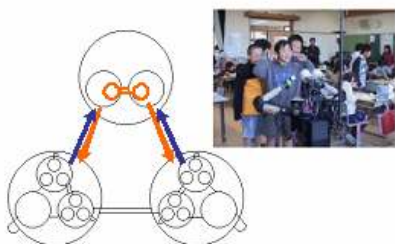


図 11: エージェントタイプ 2

4.2.3 エージェントタイプ 3

最後に, ユーザからどう思われているかの認識モジュールを持っている癒しロボットがいる場合, タイプ 3 として図 12 のように, 他者の人から好かれているか嫌われているかを, エージェントのソシオグラムにおける P モードの荷重円によって表すことができる. (5)式において $j = k$ とし, 人が人自身に対する好き嫌いを認識できるようにした.

また, タイプ 3 は自己への C モードを持っているため(13), (14)式の地位向上方略を取り入れた. これはエージェントが放置された際, 構ってほしいと鳴いたりする場合を考慮したためである.

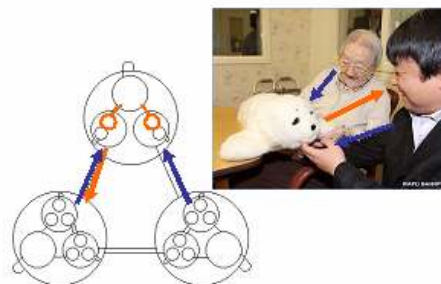


図 12: エージェントタイプ 3

今回, これらのエージェントにはバランス理論は取り入れなかった. 人間関係に左右されるのではなく, 低い好感度を与えられている人に対しても好感度を持てるようにするためである.

5 シミュレーション実験

5.1 実験内容

人の社会集団の中で, 3 節でモデル化したエージェントをコミュニケーションさせるシミュレーションを行う. シミュレーションは人だけの場合と人とタイプ 1, 人とタイプ 2, 人とタイプ 3 を比較した計 3 パターンを行った.

5.2 シミュレーションモデル

人とエージェントは以下の手順で他者とコミュニケーションをとり, P モードと好感度を更新する.

- コミュニケーション相手の選択
- 話題対象の選択
- P モードの更新
- コミュニケーション相手・話題対象の好感度更新
- 地位向上方略の実行

5.3 実験条件

本研究では, 人とエージェントのコミュニケーションの環境の集団規模として, 学校のクラス単位を考える. 文部科学省の平成 23 年度の学校基本調査人における 1 クラスの人数 29.2 人であることを踏襲し, 人のみの場合の人数は 29 人とする. そこでエージェントを参加させた場合は人が 28 人とエージェントが 1 体とする.

全ての人, エージェントにシミュレーション開始時に第一印象として他者への好感度を $[-0.2, 0.2]$ の一様分布の乱数を与える. [5]のモデルを引用し, シミュレーションは合計 2000 ターン行い, それを 100 回実行する. シミュレーションにおけるパラメータの値を表 1 に示す.

表 1: パラメータの値

η : 地位計算時の係数	1.0
α_1 : ゆらぎの大きさ 1	0.1
α_2 : ゆらぎの大きさ 2	0.2
ω_1 : 好感度更新の係数 1	0.25
ω_2 : 好感度更新の係数 2	0.2
ω_3 : 好感度更新の係数 3	1.0
χ : 好感度の差異の閾値	0.2
σ : 地位向上方略の変化量	0.1
友人閾値	0.2
排斥閾値	-0.2

5.4 評価内容

評価内容は、エージェントがいる場合といない場合の、

- ・いじめ発生確率 (グループ, クラス)
 - ・いじめを受けた人の平均人数 (グループ, クラス)
- である。いじめの発生確率はクラス内に一人以上いじめを受けている人がいる確率とした。

6 結果

いじめ発生確率といじめを受けた人の平均人数について、表 2 にまとめた。

表 2: いじめ発生確率といじめを受けた人の平均人数

	発生確率(%)		平均人数(人)	
	グループ	クラス	グループ	クラス
人	86	65	0.84	2.08
タイプ 1	37	5	0.32	0.14
タイプ 2	82	67	0.86	1.82
タイプ 3	93	59	0.91	2.00

評価内容に基づき検定を行った。人とエージェントのいじめ発生確率を図 13 に示す。エージェントタイプ 1 のいじめ発生確率について表 2 の値から検定を行った結果、有意水準 0.01 においてグループによるいじめ発生確率($0.06 \times 10^{-12} < 0.01$)、クラスによるいじめ発生確率($0.02 \times 10^{-18} < 0.01$)ともに有意差が確認された。

エージェントタイプ 2 のいじめ発生確率について、グループによるいじめ発生確率($p=0.46 > 0.1$)、クラスによるいじめ発生確率($p=0.74 > 0.1$)の有意差は確認されなかった。

エージェントタイプ 3 のいじめ発生確率について、グループによるいじめ発生確率($p=0.13 > 0.1$)、クラスによるいじめ発生確率($p=0.38 > 0.1$)の有意差は確認されなかった。

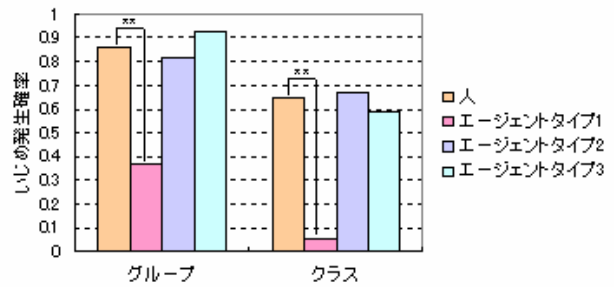


図 13: いじめ発生確率

一方、いじめを受けた人の平均人数を図 14 に示す。エージェントタイプ 1 のいじめを受けた人の平均人数について表 2 の値から検定を行った結果、有意水準 0.01 においてグループによるいじめを受けた人の平均人数($0.02 \times 10^{-7} < 0.01$)、クラスによる平均人数($0.02 \times 10^{-33} < 0.01$)ともに有意差が出た。

エージェントタイプ 2 について、グループによる平均人数($p=0.81 > 0.1$)の一方、クラスによる平均人数($p=0.042 < 0.05$)で有意水準 0.05 において有意差が見られた。

エージェントタイプ 3 について、グループによる平均人数($p=0.47 > 0.1$)とクラスによる平均人数($p=0.61 > 0.1$)双方とも有意差は見られなかった。

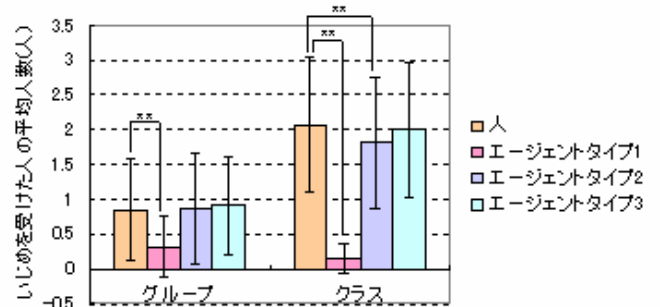


図 14: いじめを受けた人の平均人数

7 考察

ぬいぐるみや機械を表すエージェントタイプ 1 はグループ、クラスの場合ともにいじめ発生確率が下がり、いじめを受けた人の平均人数も少なくなった。また、ヒューマノイドロボットを表すエージェントタイプ 2 についてはクラスからいじめを受けた人の平均人数が少なくなることがわかった。その他のエージェントタイプについては発生確率や平均人数が増加した一方減少する場合もあったが、有意な差は得られなかった。自分の地位を向上させようと振舞っているにも関わらず、エージェントタイプ 1 のようなソシオグラムのモデルがいる場合はいじめが大

きく減少する結果となった。人からどのような好感度をもらっても、エージェントの人に対する好感度やPモードといった内部モデルが一定で変動しないため、好感度が下がる悪循環が回避されたのではないかと考える。また、エージェントタイプ2においてクラス（集団）からいじめを受けた人の平均人数が減少し有意差が出たことについては、自分に対するコミュニケーション相手からのPモードを設定していないため、相手から嫌われていたとしてもいじめであると主観的に判定していないことが一つの要因ではないかと考えた。これを考慮するとエージェントタイプ1とエージェントタイプ2は、自分 a_i の頭の中におけるコミュニケーション相手 a_j からのPモード W_{ji}^i が無いことが共通点である。そのためこのソシオグラムの有無によるいじめ発生確率や平均人数の増減を確かめ、そのソシオグラムの有無の意味も考察する必要があると考える。

8 まとめと今後の予定

本研究では、ソシオン理論を用いてモデル化したエージェントと、ソシオン理論を適用した人との関係性の変化をシミュレーションした。今後は考察の節で述べたように他のソシオグラムについてもエージェントタイプ分類を行い、シミュレーションを行う。一方でいじめが始まる時間の違いをしてみるなど評価基準の視点を増やしていく。本研究では関係性を調べるために検討したのはソシオグラムの荷重円の有無だけだったので、エージェントの違いにより人のソシオグラムを変化させ検討すると共に、人がエージェントに抱く第一印象などを考えていきたい。

参考文献

- [1] T.Kanda, R.Sato, N.Saiwaki and H.Ishiguro: Atwo-month field trial in an elementary school for long-term human-robot interaction, IEEE Transaction on Robotics, Vol. 23, No. 5, pp. 962-971, (2007)
- [2] 宇津木到, 三上達也: 局所的人間関係におけるマルチエージェントシミュレーションの研究, 政策科学, Vol. 14, No. 2, pp. 171-184, (2007)
- [3] 坂本大介, 小野哲雄: ロボットの社会性: ロボットが対話者間の印象形成に与える影響評価, ヒューマンインターフェース学会論文誌, Vol. 8, No. 3, pp. 381-390, (2006)
- [4] 木村洋二, 藤澤等, 雨宮俊彦: ソシオンの理論(1), 関西大学社会学部紀要, vol21(2), pp.67-143, (1990)
- [5] 大隈俊宏, 大澤博隆, 今井倫太: ソシオン理論に基づいた学校のクラス内のいじめのモデル化, Joint Agent

Workshop and Symposium(JAWS), (採録決定済み), (2012)

- [6] 藤澤等: ソシオン理論のコア, 北大路書房, (1997)
- [7] 田中恵海, 高橋謙輔, 鳥海不二夫, 菅原俊治: 学級のいじめ問題を題材とする工学的シミュレーションとその考察, 情報処理学会論文誌, 数理モデル化と応用 3(1), pp.98-108, (2010)