ユーザの報酬付与頻度が ロボットとのインタラクションに与える影響 – ボールを使ったやりとり遊びの学習と印象評価 –

User's Reward Frequency Influencing on Impression of Robot During Interaction: learning and impression evaluation of ball play

佐久間 拓人 [1*] 加藤 昇平 [1]
Takuto Sakuma[1] Shohei Kato[1]

[1] 名古屋工業大学 大学院工学研究科 情報工学専攻
[1] Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

Abstract: We focused our attention on human-robot interaction for one purpose that users can really enjoy it. We believed that it is necessary to reflect the preference of users so that they could have more positive impressions of robot through such interactions. This research was aimed at developing a robot that can reflect the preference of users. In this robot, users reward the interaction between robot and themselves, and then the robot learns reward dynamically. By doing that the robot can create better interactions for users. Thereby, users' impression of robot can be improved. In order to assess the effectiveness of our robot, we did an experiment that applied a turn-taking play using a ball as the interaction between user and robot. And according to sensitivity evaluation our robot gained the best impression of users.

1 はじめに

人とロボットの間におけるインタラクションは技術の発展に伴い，多種多様に発展してきた．最近では人型ロボットも増え，QRIO[Tanaka 05] や Robovie[Mitsunaga 06] など人とのインタラクションを目的としたロボットも数多く開発され，そのようなロボットとのインタラクションを題材とした研究もまた数多く見られる．

インタラクションを目的としたロボットの多くはユーザとのインタラクションの継続のため，「人を飽きさせない」ように行動指針を設けられていると考える．はじめは興味深々でもインタラクションが単調であれば，人が次第に飽きてゆき，やがてロボットとのインタラクションそのものを行わなくなってしまう．この問題に対し，栗山らは子どもを模したロボットを相手にやりとり遊びをし，やりとりルールを共創するしくみを提案している [栗山 10]．栗山らは人間同士のやりとりにみられる性質に着目し，人間同士のようにやりとりを通じてやりとりを広げ，共有感を持ちながら，飽きずに長く付き合っていけるロボットを目指している．

*連絡先：名古屋工業大学 大学院工学研究科 情報工学専攻 加藤昇平研究室
〒 466-8555 愛知県名古屋市昭和区御器所町
E-mail: shohey@katolab.nitech.ac.jp

我々はインタラクションにおける重要な要素として「相手に与える印象」があると考える．インタラクションを行う相手が不快な印象を抱けば，インタラクションに支障をきたす可能性は高い．逆に相手が良い印象を抱くことによって，例え稚拙なインタラクションであっても継続して行われる可能性は高いと考える．阿部らは子どもの心理状態を推定し，適切な行動を選択することで子どもを飽きさせず長い間インタラクションを続けることが出来る遊び相手ロボットモデルを構築している [阿部 11]．しかし，これらの研究の多くはあくまで人の心理状態を推定しており，扱うインタラクションによって推定モデルを変更する必要がある．

本研究はユーザからの報酬を正確に学習し，インタラクションにユーザの好みを反映することでユーザのロボットに対する印象を向上させることを目的としている．そのため，ユーザの報酬は推定ではなく，ユーザに明示的に与えさせることとした．ユーザとロボットはインタラクションを行い，ユーザは自分の好みに従いインタラクションに報酬を付与する．ロボットは報酬を基にユーザの好みを学習し，次のインタラクションに反映する．

本稿では先行研究 [佐久間 14a] にて行った「ボールを使ったやりとり遊び」に対する感性評価実験の結果

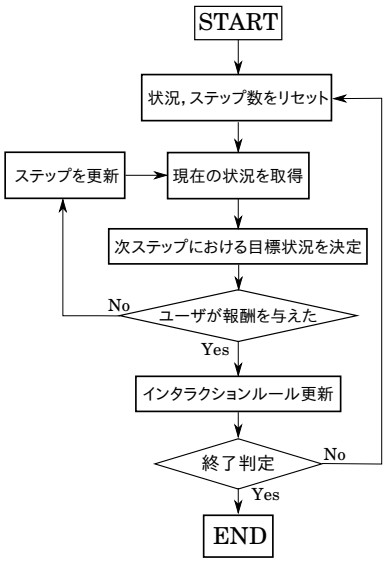

図 1: ロボットの行動決定プロセス

を基に，提案手法の有効性を報酬付与頻度の違いに着目し検証する．

2 ユーザの報酬付与傾向の獲得

2.1 インタラクションの流れ

本稿ではユーザ報酬を取り入れたインタラクションモデルを使用する．図 1 にユーザ報酬を取り入れたインタラクションを行うロボットの行動決定プロセスを示す．なお，本稿ではユーザとロボットの姿勢（手先位置）やボールの位置情報などインタラクションにかかわる情報を「状況」と呼び，ロボットは 1 ステップ毎に状況を取得可能であるとする． ロボットは直前のステップまでの状況とインタラクションルールから次ステップにおける目標状況を決定する．インタラクションルールの更新はユーザ報酬を基に行われる．ユーザから報酬付与が行われた場合，状況，ステップ数をリセットし初期状況からやりとりを再開するとする．

2.2 インタラクションルールの更新

ユーザが与えた報酬をインタラクション系列全体への報酬としてのみ捉えると，報酬付与時にユーザがどの状況に注目して報酬を与えたのか，どのような意図で報酬を与えたのかをロボットは把握出来ない．よって，ユーザの報酬付与傾向を詳細に獲得する手法として，All-Combinatorial N-gram（ACN）[佐久間 14b] を導入する．本稿では ACN の分割によって生成されたパターン系列の集合をインタラクションルール $\boldsymbol{R}$ と呼ぶ．$\boldsymbol{R}$ の概要を図 2 に示す．$\boldsymbol{R}$ は今までの経験を保持しており，ユーザが報酬を与える度に更新される．また，$\boldsymbol{R}$ の更新とは N-gram ルール内の各パターンルールの報酬履歴を更新することとする．ここでパターンルール R_{ij} とは，ACN によって分割されたパターン系列 D_{ij} と D_{ij} の過去の報酬履歴を格納しているものとする．N-gram ルールはパターンルールの長さ毎の集合である．パターン系列 D_{ij} のある報酬値 P_{DL} はパターン系列の長さおよび報酬付与時からの近さで重みを算出される．式 (1)，(2) にそれぞれパターン系列の長さによって報酬値 P_L を決定する式，報酬付与時からの近さによって報酬値 P_D を決定する式を示す．

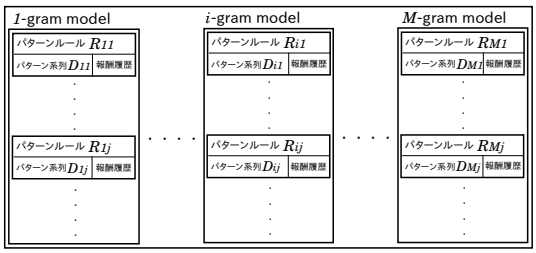

図 2: インタラクションルール概要

$$P_L = \frac{U \times A_{ij}}{(1 + e^{T-||D_{ij}||-T_{max}/2})}. \quad (1)$$

$$P_D = \frac{U \times A_{ij}}{(1 + e^{T-K_{ij}-T_{max}/2})}. \quad (2)$$

ここで i は N の値を表し，j は i-gram モデル内の D の識別子，U はユーザが与えた報酬を表す値，A_{ij} は D_{ij} がインタラクション系列に出現した回数，T はインタラクション系列長，$||D_{ij}||$ は D_{ij} のパターン系列長，T_{max} は最大インタラクション系列長，K_{ij} は報酬付与時から D_{ij} の末尾までの距離（記号数）を表す．本稿では U（$-X \leq U \leq X$）は整数であり，U の値が大きいほど良い報酬を意味する．なお X は任意の非負整数であり，本稿における具体的な値および報酬付与方法は 4.1 章にて述べる．式 (1) によってパターン系列 D_{ij} の報酬値 P_L を決定することでより長い系列，すなわちやりとり全体の流れに重きをおいた学習を行うことができる．同様に式 (2) によってパターン系列 D_{ij} の報酬値 P_D を決定することでより報酬付与時に近い系列，すなわち報酬タイミングに重きをおいた学習を行うことができる．本稿では式 (1)，(2) を複合し，長さによる重み，近さによる重みの双方を考慮した式によって報酬値 P_{DL} を算出する．式 (3) に本稿で使用する報酬値 P_{DL} の算出式を示す．

$$P_{DL} = \frac{U \times A_{ij}}{(1 + e^{T-||D_{ij}||-T_{max}/2})(1 + e^{T-K_{ij}-T_{max}/2})}. \quad (3)$$

3　目標状況の決定

ロボットは目標状況を決定したのち，その状況に移行するための行動を出力する．次ステップにおけるロボットの目標状況は獲得したインタラクションルール $\boldsymbol{R}$ 及び，直前のステップまでのパターン系列によって決定される．ロボットはユーザからの報酬を最大化するような状況を目標とする．

ロボットが選択可能な状況の集合を $\boldsymbol{\alpha}$ とし，目標状況の候補を $\alpha_k \in \boldsymbol{\alpha}$ とすると，ロボットは α_k に対する総報酬予測値 E_{α_k} を以下の手順で決定する．

1. パターン系列に α_k を含むパターンルールの内，α_k 以前の系列が存在しない，あるいはやりとりにおける直前のステップまでの状況の系列と一致するパターンルール，および後方一致するパターンルールの集合 $\boldsymbol{R}_{\{\alpha_k\}}$ を求める．

2. α_k の総報酬予測値 E_{α_k} を下式にて求める．

$$E_{\alpha_k} = \sum_{R \in \boldsymbol{R}_{\{\alpha_k\}}} F(R). \tag{4}$$

$$F(R) = \begin{cases} \mu(R) \times \frac{1}{\sqrt{2\pi\sigma(R)}} & (\sigma(R) \neq 0), \\ \mu(R) & (\sigma(R) = 0). \end{cases} \tag{5}$$

ここで $\mu(R)$ は R が持つ報酬履歴に含まれる報酬値 P_* の平均値を，$\sigma(R)$ は標準偏差を表す．なお P_* は報酬値 P_L, P_D, P_{DL} のいずれかを表す．

以上の手順を状況集合 $\boldsymbol{\alpha}$ 内の要素全てに対して行い，算出された目標状況の候補 α_k の総報酬期待値 E_{α_k} から相対的に α_k の生起確率を算出，確率的に目標状況を決定する．

これにより決定した目標状況に基づき行動を出力することでユーザから高報酬が得られる確率の高い行動となる．なお，抽出されたパターンルールが一つも無い場合（$R_{\{\alpha_k\}} = \phi$ for $\forall \alpha_k \in \boldsymbol{\alpha}$）はランダムに目標状況を決定する．

4　感性評価実験

人－ロボット間のインタラクションにおいて本稿で提案した学習手法の有効性を確認するため，感性評価実験を行った．

4.1　計算機実装

本稿では人とロボットとのインタラクションに使用するインターフェースとして栗山ら [栗山 10] が使用した環境を参考に同様の環境を用意した．すなわち，ボールを使った遊びに場面を設定し，シミュレータ上に擬人化エージェントの体を二体構築し，向き合う形で被験者が操作するエージェントとロボットが操作するエージェントを配置した．なお本稿では被験者が操作するエージェントを「ユーザエージェント」，ロボットが操作するエージェントを「ロボットエージェント」と呼称する．ユーザエージェントとロボットエージェントの間にはボールを配置した．図 3 に実験に用いた擬人化エージェントおよびボールの外観を示す．紙面の都合上，インターフェースの詳細は先行研究 [佐久間 14a] を参照して頂きたい．

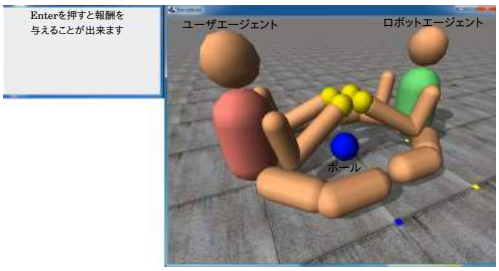

図 3: インタラクションに用いた擬人化エージェント

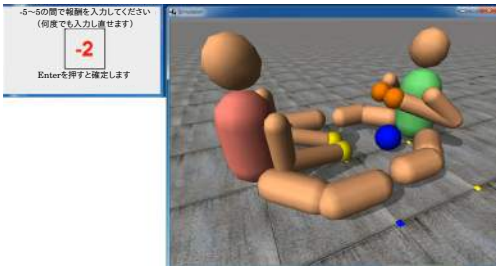

図 4: 報酬付与中の様子

図 4 に被験者がロボットエージェントに報酬を与えている様子を示す．本稿ではロボットエージェントとのインタラクションについて「良いやりとり」であると感じた場合に高い報酬を与えるよう被験者に指示する．どのようなやりとりを「良い」と思うかは被験者の主観に任せるものとする．被験者からの報酬値 U は-5（とても悪い）から+5（とても良い）の間の 11 段階（$X = 5$）とする．

4.2　感性評価

被験者には 5 つのロボットとやりとりさせ，各ロボットとのやりとり終了後に感性評価をしてもらった．評価実験に用いたロボットを以下に示す．

- 提案ロボット（ロボット DL）：
  提案手法によりユーザ報酬付与傾向を学習するロボット

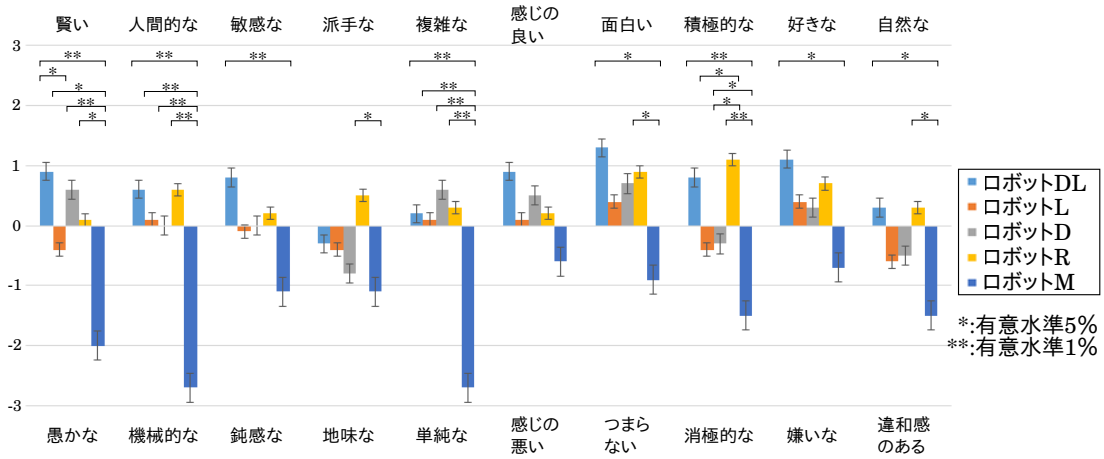

図 5: 感性評価実験の結果

- 長さ優先ロボット（ロボット L）：
  提案手法のうちインタラクションルール更新時に式 (1) を使用するロボット
- 近さ優先ロボット（ロボット D）：
  提案手法のうちインタラクションルール更新時に式 (2) を使用するロボット
- ランダムロボット（ロボット R）：
  ユーザからの報酬を無視し，ランダムに動くロボット
- ミラーリングロボット（ロボット M）：
  ユーザの動きを真似するロボット

なお，長さ優先ロボットおよび近さ優先ロボットは本稿で提案したインタラクションルール更新時に用いる重み（式 (3)）の有効性の検証のため採用した．それぞれ更新時に式 (1)（計算時に系列の長さのみを考慮する式）を用いるロボット，式 (2)（報酬付与時からの近さのみを考慮する式）を用いるロボットとなっている．また，ミラーリングロボットはロボットの状況取得時におけるユーザエージェントの手の状態と同じ状態になるよう出力を決定する．感性評価には SD 法を用い，各形容詞対（図 5 参照）について 7 段階評価で行う．

また，全ロボットとのインタラクション後に「最も印象が良かったロボット」および「最も印象が悪かったロボット」を聴取した．

図 5 に感性評価実験の結果を示す．棒グラフはユーザの感性評価の平均を，誤差棒は標準誤差を表す．また各ロボットの評価に対して Tukey の多重比較検定による有意差検定を行った．検定の結果，有意水準 1 %および 5 %にてロボット間に有意差を確認できたものを「＊＊」および「＊」でそれぞれ示す．

図 5 より提案ロボット（ロボット DL）は多くの形容詞対でミラーリングロボット（ロボット M）に対して有意差を確認できたことが分かる．ランダムロボット（ロボット R）に対しては「賢い」,「敏感な」,「感じの良い」,「面白い」,「好きな」という評価において提案ロボットの方が高いことを確認した．しかし，有意差検定における有意差は確認できなかった．ミラーリングロボットは全形容詞対に対してネガティブな評価をされており，本稿におけるインタラクションには向いていないロボットであったと考える．我々が以前行ったインタラクション実験 [佐久間 14b] においてはミラーリングロボットが必ずネガティブな評価をされることは稀であり，扱うインタラクションによって有効性が大きく変わるロボットであると考えられる．

表 1: 最も印象が良かった・悪かったロボット

被験者	最も良い	最も悪い
A	**ロボット DL**	ロボット M
B	ロボット D	ロボット M
C	ロボット D	ロボット M
D	**ロボット DL**	ロボット M
E	ロボット M	ロボット R
F	ロボット D	ロボット R
G	ロボット R	ロボット M
H	**ロボット DL**	ロボット M
I	**ロボット DL**	ロボット D
J	**ロボット DL**	ロボット D

表 1 に被験者が「最も印象が良かった，または悪かったロボット」というアンケートに回答した結果を示す．表 1 より，提案ロボット（ロボット DL）が最も多く「最も印象が良かった」と評価されたことがわかる．これにより，感性評価実験ではランダムロボット（ロボット R）との有意な差は確認出来なかったものの，被験者の中では「最も印象が良かった」ロボットである傾向がみられることから提案ロボットの有効性が示せたと考える．なお，近さ優先ロボット（ロボット D) は提案ロボットについで多く「最も印象が良かった」と評価されたが,「最も印象が悪かった」とも評価されており，被験者によっては報酬付与時に近い系列であるほど報酬への影響が大きいとは限らないと考えられる．また，ミラーリングロボット（ロボット M）が最も多く「最も印象が悪かった」と評価されており，感性評価実験

表 2: 被験者の報酬付与頻度

被験者	報酬付与頻度	群	平均報酬付与間隔（備考）
A	0.96	消極	40.75
B	2.92	積極	15.93
C	0.68	消極	55.36
D	1.52	積極	22.31
E	1.28	消極	34.89
F	0.88	消極	49.17
G	1.92	積極	20.72
H	1.12	消極	36.28
I	2.08	積極	20.06
J	0.44	消極	78.60
平均	1.38		37.41

の結果と同等の結果であることを確認した．

5 報酬付与頻度の影響

本稿で構築したインターフェースではユーザは好きなタイミングでロボットエージェントに報酬を与えることが出来る．そのため被験者ごとに報酬を与えるタイミングや回数が異なる．表 2 に感性評価実験において各被験者がロボットに与えた報酬の頻度[1]を示す．報酬付与頻度が被験者平均よりも高い被験者は積極的にロボットに報酬を与えていると考えられるため「積極群」に分類し，報酬付与頻度が被験者平均以下の被験者は「消極群」に分類する．

5.1 感性評価の違い

図 6，図 7 に群ごとの感性評価の結果を示す．棒グラフはユーザの感性評価の平均を，誤差棒は標準誤差を表し，有意水準 1 %および 5 %にてロボット間に有意差を確認できたものを「＊＊」および「＊」でそれぞれ示す．図 6 と 7 より，積極群は消極群に比べ全体的にポジティブな評価に偏っていることが分かる．特に「派手な」「自然な」の形容詞対に関しては消極群に対して積極群の評価は全ロボットで上昇している．なお，両形容詞対に関しては Mann-Whitney の U 検定により分布に差があることも確認した（有意水準 5 %）．ロボット間の差に着目すると，長さ優先ロボット（ロボット L）の評価に著しい差が見て取れる．特に「派手な」「積極的な」の形容詞対に関しては Tukey の多重比較検定により有意差を確認した（有意水準 5 %）．本実験における全てのロボットは学習できる系列の長さに制限がある（$T_{max} = 8$）．すなわちインタラクションルール更新に用いられる系列はユーザが報酬を与えた時点からさかのぼって最新 T_{max} 個までである．一方で，消極群の被験者は報酬付与頻度が低いため報酬の間隔は長くなる．本実験の場合，消極群に属する全ての被験者において報酬付与間隔は全て T_{max} を超えており，常に $T = T_{max}$ となるため式 (1) の影響が低くなると考える．逆に積極群においては報酬の間隔が短いために更新時の系列も短くなり，式 (1) の影響が大きくなると考えられる．これによりロボット L に対する群間の差が生じたのではないかと考える．

[1]各ロボットとのインタラクションにおける 1 分間当たりの報酬付与回数の平均値

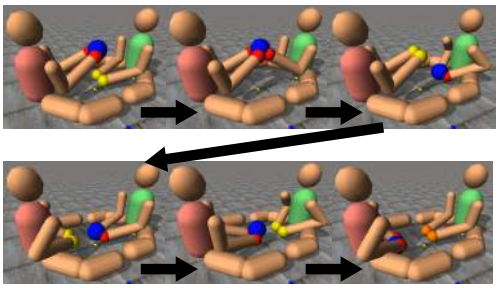

図 8: 「ボールの受け渡し」の様子

またランダムロボット（ロボット R）に着目すると，積極群においてランダムロボットは最も評価が高い．しかし消極群においては「派手な」「積極的な」の形容詞対以外の評価が低くなっている．この群間の差は観測するインタラクションの長さの違いが影響していると考える．積極群は報酬付与頻度が高く，報酬の間隔が短い．すなわち報酬と報酬の間で行われるインタラクションは短く，例えランダムに動いていてもそれをランダムと認知し難いと考える．逆に消極群は観測するインタラクションが長く，ロボット R の動きに学習している様子が見られないことを認知しやすいのではないかと考える．これは「賢い」の形容詞対におけるロボット R の評価が，積極群と消極群で大きく違うことからも示唆される．

ここで着目すべきは何故積極群において提案ロボット（ロボット DL）の評価が低いのか．本来積極的に報酬を与えられた方が提案ロボットはより学習し，より良いインタラクションを行うことが出来るはずである．その原因を本実験において実際に創発されたやりとりを例に考察していく．

5.2 やりとりの違い

ここで，各被験者群で顕著に見られたやりとりの様子を述べる．積極群はユーザエージェントが差し出したボールをロボットエージェントが受け取ったら報酬を与える，あるいはロボットエージェントが地面に落ちているボールを拾い，ユーザエージェントに向かって差し出してきたら報酬を与えるなど，一方通行のやりとりが成立したときに報酬を与える様子が見られた．それに比べ消極群は，積極群に比べ長い間やりとりを続け，主に双方向のやりとりが成立した時に報酬を与

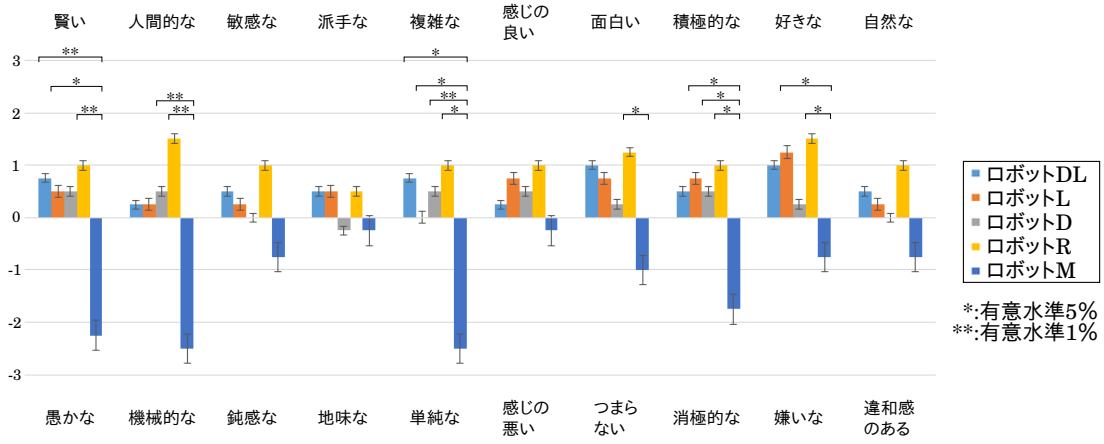

図 6: 積極群の感性評価実験の結果

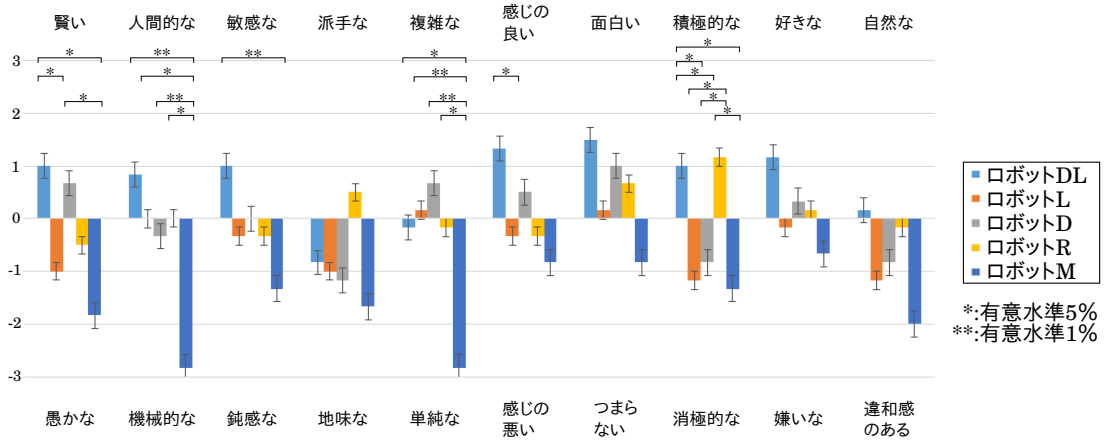

図 7: 消極群の感性評価実験の結果

える様子が見られた．このように一方通行のやりとりは積極群で多く，双方向のやりとりは消極群で多くみられたことが，感性評価へ影響を与えたと考える．特に双方向のやりとりは通常一方通行のやりとりよりも成立しづらいため，成立したとき被験者がポジティブな印象を持つことが考えられる．

典型的な双方向のやりとり例を図 8 に示す．図 8 はボールをユーザエージェントが持ち，そのボールをロボットエージェントが受け取り．その後ロボットエージェントがユーザエージェントにボールを渡すやりとりを示している．これを「ボールの受け渡し」と呼ぶ．このやりとりは本実験で行われた全 50 回のやりとり中，14 回のやりとりで成立が確認された．「ボールの受け渡し」を成立させたやりとりを行なっていた被験者は 10 名中 6 名であり，特に被験者 F はミラーリングロボット以外のロボット全てにおいてこのやりとりを成立させていた．これは被験者 F が消極群であったためと考える．なお,「ボールの受け渡し」はボールを使った

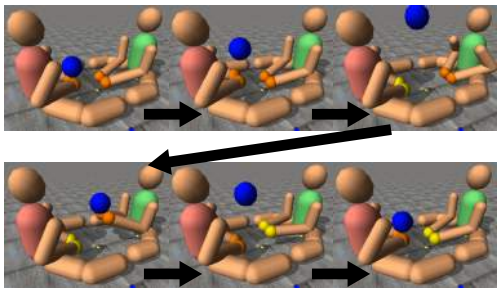

図 9: 「バレーボールのようなやりとり」の様子

やりとり遊びとして一般的であり比較的成立しやすいため，このやりとりの成立がそのまま感性評価にポジティブな影響を与えている被験者は 2 名のみであった．

成立しづらい双方向のやりとり例を図 9 に示す．図 9 はボールをユーザエージェントが握った状態の手で上方向に飛ばし，それをロボットエージェントが握っ

表 3: ロボットごとの報酬付与頻度

	報酬付与頻度（全被験者平均）
提案ロボット	**1.72**
長さ優先ロボット	1.3
近さ優先ロボット	1.42
ランダムロボット	1.58
ミラーリングロボット	**0.88**

た状態の手で弾き，ユーザエージェントが握った状態の手で受け止めるやりとりを示している．これを「バレーボールのようなやりとり」と呼ぶ．このやりとりはボールの物理的制約を考慮した上で行われる極めて難しいインタラクションであると我々は考える．本実験で行われた全 50 回のやりとり中，成立が確認されたのは 6 回であった．この「バレーボールのようなやりとり」は被験者 C と J においてよく観測された．共に消極群の被験者である．特に被験者 C は「バレーボールのようなやりとり」の成立を狙って報酬を与えていたと実験後のインタビューで答えており，実際にこのやりとりが成立したロボット D への感性評価は非常にポジティブであった．

5.3 まとめ

以上の結果から被験者ごとに異なる報酬付与頻度がインタラクションおよび感性評価に対して大きな影響を持っていることを確認した．

ではなぜ，被験者ごとに報酬付与頻度が大きく異なるのか．我々は被験者がインタラクション相手のロボットに対して学習することを期待しているかどうかによるのではないかと考えた．我々は感性評価実験に用いたロボットごとの報酬付与頻度の違いに着目した．表 3 にロボットごとの報酬付与頻度における全被験者平均を示す．表 3 より，ミラーリングロボットに対する報酬付与頻度が最も低いことが分かる．これはミラーリングロボットに対して学習を期待していないからではないかと考える．ミラーリングロボットの「ユーザの動きを真似する」仕組みには全被験者がインタラクション中に気づいており，ミラーリングロボットが学習することに期待を持てなくなったと考えられる．なお，提案ロボットに対する報酬付与頻度は最も高かったため，提案ロボットに対しては学習を期待する被験者が多かったことが示唆される．

今後は報酬付与頻度を固定した上での実験なども行い，最適な報酬付与頻度やユーザの報酬付与頻度に合わせた学習を行うロボット，あるいは表情やジェスチャーなどによってロボットの学習度合いを示すことで，ユーザに報酬付与を促すロボットを考案していく．

6 おわりに

本稿では先行研究 [佐久間 14a] にて行った実験に対して報酬付与頻度という新たな着眼点を基に考察，検証を行った．報酬付与頻度の高い被験者群と低い被験者群の結果を比較し，提案手法の有効性および今後検討すべき項目を明らかにした．今後は検証をさらに進め，ロボットの選択できる行動の種類を拡大した複雑なインタラクションを扱えるロボットを構築するとともに，既存の機械学習手法などとの性能比較を行なっていく．

参考文献

[阿部 11] 阿部 香澄, 岩崎 安希子, 中村 友昭, 長井 隆行, 横山 絢美, 下斗米 貴之, 岡田 浩之, 大森 隆司：子供と遊ぶロボット:他者の状態推定に基づく行動決定モデルの適用, HAI シンポジウム, pp. I–2B–3 (2011)

[栗山 10] 栗山 貴嗣, 國吉 康夫：応答予測と馴化・脱馴化に基づき人とやりとりルールを探索・共創するロボットモデル, 日本ロボット学会誌, Vol. 28, No. 8, pp. 1036–1046 (2010)

[Mitsunaga 06] Mitsunaga, N., Miyashita, T., Ishiguro, H., Kogure, K., and Hagita, N.: Robovie-IV: A Communication Robot Interacting with People Daily in an Office, in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pp. 5066–5072 (2006)

[佐久間 14a] 佐久間 拓人, 加藤 昇平：All-Combinatorial N-gram に基づく擬人化エージェントによるボールを使ったやりとり遊び, 第 175 回情報処理学会 知能システム研究会, pp. No.13 (6–pages) (2014)

[佐久間 14b] 佐久間 拓人, 加藤 昇平：ユーザ評価傾向の動的獲得によるヒューマンインタラクションの創発, 電気学会論文誌, Vol. 134-C, No. 2, pp. 303–311 (2014)

[Tanaka 05] Tanaka, F., Fortenberry, B., Aisaka, K., and Movellan, J. R.: Developing dance interaction between QRIO and toddlers in a classroom environment: plans for the first steps, in *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pp. 223–228 (2005)