1E5-OS-23b-6

インタラクションにおけるユーザ評価傾向獲得 –ボール遊びを例として–

Acquiring User's Evaluation Tendency in Interaction: the case of ball play

佐久間 拓人　　加藤 昇平
Sakuma Takuto　　Kato Shohei

名古屋工業大学工学研究科情報工学専攻
Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

We paid our attention to the interaction for purpose of enjoying turn-taking. We believe that it is necessary to reflect the tendency of the user for an interaction so that a user has a positive impression for the system through such an interaction more. This research aims at development of the human interaction system which reflect a tendency of the user. The user gives an evaluation to the user and the system's interaction, which the system learns dynamically. The system creates a better interaction for the user. Thereby, the impression of the system from the user improves. This report pays attention to turn-taking play using a ball, and inspects the effectiveness of proposal method in a subjective evaluation experiment.

1. はじめに

近年，Human-Robot Interaction(HRI) に関する研究が盛んに行われている [Mitsunaga 06, Osada 06]．ロボット技術が人間の生活空間に浸透するのに伴い，人間がロボットとインタラクションを行う機会も増えていくと予想される．こうしたロボットが人間に対してどのような印象を与えるのか，人間がどのような印象を受け取るのかという点については，ロボットやインタラクションの設計において非常に重要な課題である．

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．

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．

本研究はユーザの評価を正確に学習し，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，ユーザの評価は推定ではなく，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を評価する．ロボットは評価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．

連絡先: 加藤昇平，名古屋工業大学，愛知県名古屋市昭和区御器所町，052-735-5625，shohey@katolab.nitech.ac.jp

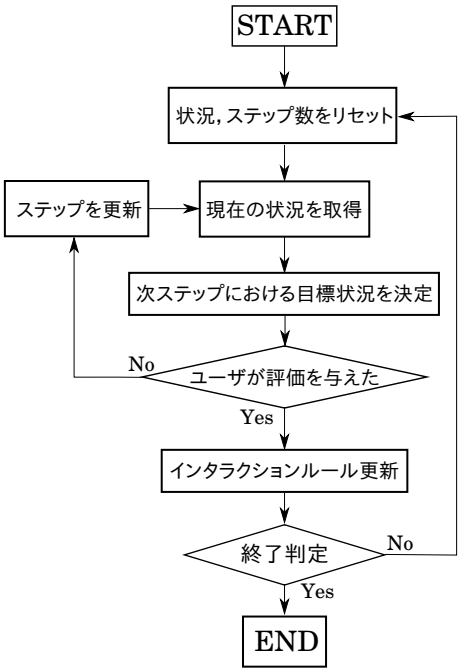

図 1: ロボットの行動決定プロセス

本稿では「ボールを使ったやりとり遊び」における学習モデルの提案，および提案手法を用いたインタラクション実験を行い，創発されたインタラクションおよび被験者による感性評価を基に有効性を示す．

2. ユーザ評価傾向の獲得

2.1 インタラクションの流れ

本稿ではユーザ評価を取り入れたインタラクションモデルを使用する．図 1 にユーザ評価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

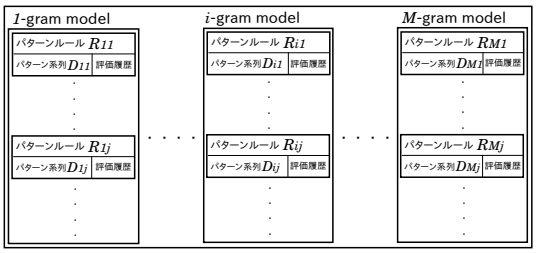

図 2: インタラクションルール概要

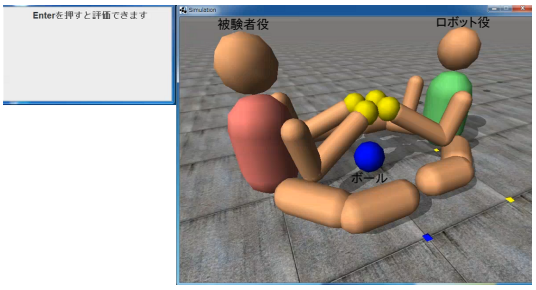

図 3: インタラクションに用いた擬人化エージェント

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評価を基に行われる．ユーザ評価が行われた場合，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．初期状況からユーザが評価を与えるまでの状況の系列をインタラクション系列と呼ぶ．

2.2 インタラクションルールの更新

ユーザが与えた評価をインタラクション系列全体への評価としてのみ捉えると，評価時にユーザがどの状況に注目して評価を与えたのか，どのような意図で評価を与えたのかをロボットは把握出来ない．よって，ユーザの評価傾向を詳細に獲得する手法として，All-Combinatorial N-gram（ACN）[佐久間 14] 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，ユーザが評価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評価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} の過去の評価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} のある評価値 P はパターン系列の長さおよび評価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評価値 P を決定する式，評価時からの近さによって評価値 P を決定する式を示す．

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

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

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 はユーザが与えた評価を表す値，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} は評価時から 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評価を意味する．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評価方法は 4.1 章にて述べる．式 (1) によってパターン系列 D_{ij} の評価値 P 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} の評価値 P を決定することでより評価時に近い系列，すなわち評価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評価値 P を算出する．式 (3) に本稿で使用する評価値 P の算出式を示す．

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

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評価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 に対する総評価予測値 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 の総評価予測値 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 が持つ評価履歴に含まれる評価値の平均値を，$\sigma(R)$ は標準偏差を表す．

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 の総評価期待値 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評価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，感性評価実験を行う．

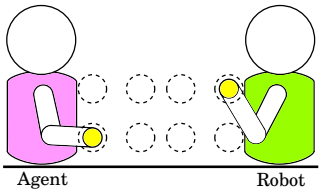

図 4: 手の位置

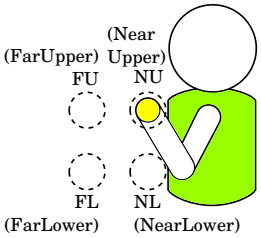

図 5: 手を移動できる場所

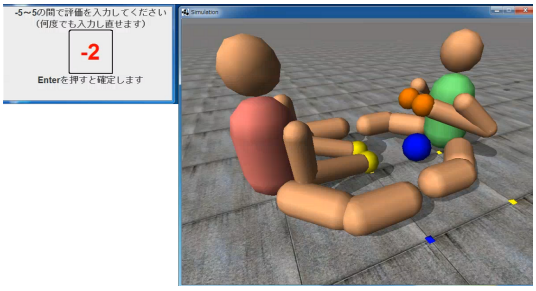

図 6: 評価付与中の様子

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システム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「被験者役」，システムが操作するエージェントを「ロボット役」と呼称する．被験者役とロボット役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．

被験者役は被験者がキーボードで特定のボタンを押すことで操作する．簡単のため，各エージェントの手とボールの位置は矢状面（両者にとって上下前後の面）に拘束する．ロボット役と被験者役は手の位置，手の握りを変えることができ，手を握っていないときは手は黄色に表示，手を握ってボールを掴んだ場合は赤く表示，手を握ってボールが手に触れていなかった場合はオレンジ色に表示される．また，物理現象をインタラクションに利用できるようにするため，動力学シミュレータである ODE（Open Dynamics Engine）を導入した．

本稿では被験者役とロボット役の手の位置の状態数は 4 であるとする．すなわち，両者の手先位置は予め定めた 4 箇所（図 4）にのみ移動できるものとする．また 4 箇所の名称はそれぞれ，NearUpper（NU），NearLower（NL），FarUpper（FU），FarLower（FL）とする（図 5）．なお，本稿における「手の位置」とは「両手の位置」である．すなわち，両者の左右の手は拘束されており，左右の手それぞれが異なる状態に移動することは無いとする．手の握りは「握っている」,「握っていない」の 2 状態とする．ロボット役が選択できる行動の種類は {NU, NL，FU，FL} の手の位置 4 つそれぞれに対し手を握っているかどうかの 2 値を考慮した 8 種類とする．ボールの状態数は「被験者役がボールを掴んでいる」,「ロボット役がボールを掴んでいる」,「誰も掴んでいない」,「どちらも掴んでいる」の 4 状態とした．ボールは手がボールに触れている時に手を握ると掴むことが出来る．システムは 1 秒毎に状況を取得する．取得する状況に含まれる情報は，被験者役の手の状態，ロボット役の手の状態，ボールの状態の 3 次元とした．システムは状況を取得後，1 秒後の目標位置を決定し，目標位置を達成するようロボット役の手を操作する．システムはインタラクション中，取得した状況を最新 8（$T_{max} = 8$）個を記憶する．よってインタラクション系列の最大長は 8 となり，被験者から与えられた評価は最新 8 個の系列に対して与えられる．なお，やりとり開始時は何も記憶していないとする．

図 6 に被験者がロボット役に評価を与えている様子を示す．本稿ではロボット役とのインタラクションについて「良いパターン」であると感じた場合に高い評価を与えるよう被験者に指示する．どのようなパターンを「良い」と思うかは被験者の主観に任せるものとする．被験者からの評価値 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 つのシステムとやりとりさせ，各システムとのやりとり終了後に感性評価をしてもらった．評価実験に用いたシステムを以下に示す．

- 提案システム（システム DL）：
  提案手法によりユーザ評価傾向を学習したシステム
- 長さ優先システム（システム 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) を使用するシステム
- 近さ優先システム（システム 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) を使用するシステム
- ランダムシステム（システム R）：
  ユーザの出力によらずランダムに出力するシステム
- ミラーリングシステム（システム M）：
  ユーザの出力をそのまま返すシステム

なお，長さ優先システムおよび近さ優先システム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）を用いるシステム，式 (2)（評価時からの近さのみを考慮）を用いるシステムとなっている．また，ミラーリングシステムはロボットの状況取得時における被験者役の手の状態と同じ状態になるよう出力を決定する．すなわち，状況取得時の被験者役の手の状況が「NU・握っていない」であれば，ミラーリングシステムは「NU・握っていない」を目標状況とし行動を出力する．感性評価には SD 法を用い，各形容詞対（図 7 参照）について 7 段階評価で行う．

図 7 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．また各システムの評価に対して 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 %にてシステム間に有意差を確認できたものを「*」で示す．

図 7 より提案システム（システム DL）は多くの形容詞対でミラーリングシステム（システム M）に対して有意差を確認できたことが分かる．ランダムシステム（システム R）に対しては「賢い」「敏感な」「感じの良い」「面白い」「好きな」の印象において提案システムの方がポジティブに評価されたことを確認した．しかし，有意差は確認できなかった．ミラーリングシステムは全形容詞対に対してネガティブな評価がされており，本稿におけるインタラクションには向いていないシステムであったと考える．我々が以前行ったインタラクション実験

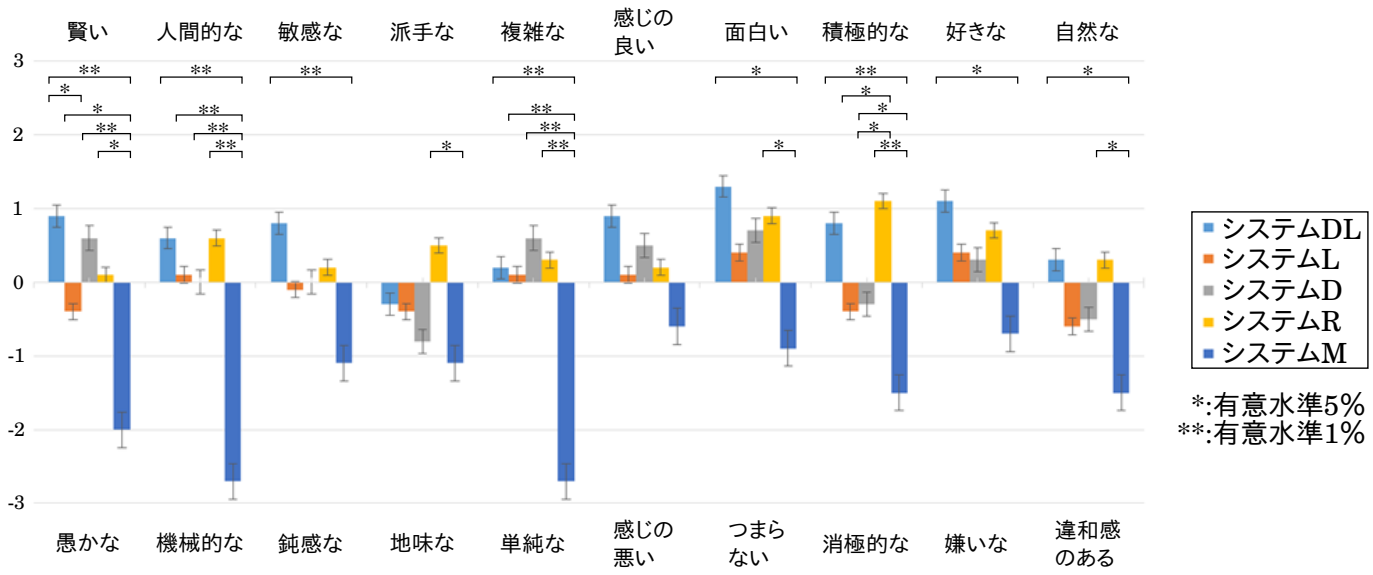

図 7: 感性評価実験の結果

表 1: 最も印象が良かった・悪かったシステム

被験者	最も良い	最も悪い
A	**システム DL**	システム M
B	システム D	システム M
C	システム D	システム M
D	**システム DL**	システム M
E	システム M	システム R
F	システム D	システム R
G	システム R	システム M
H	**システム DL**	システム M
J	**システム DL**	システム D
K	**システム DL**	システム D

[佐久間 14] においてはミラーリングシステムがネガティブな評価をされることは稀であった．本稿におけるインタラクションは以前行ったものよりも多様的であり，ミラーリングシステムのような単純なシステムではその多様性を生かしきれないのではないかと考える．

表 1 に被験者が「最も印象が良かった（または悪かった）システム」というアンケートに回答した結果を示す．表 1 より，提案システム（システム 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システム（システム 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」システムである傾向が確認された．なお，近さ優先システム（システム D) は提案システムについで多く「最も印象が良かった」と評価されたが，一部の被験者から「最も印象が悪かった」とも評価されており，被験者によっては評価時に近い系列であるほど評価への影響が大きいとは限らないと考えられる．また，ミラーリングシステム（システム M）が最も多く「最も印象が悪かった」と評価されており，感性評価実験の結果と同等の結果であることを確認した．

5. おわりに

本稿では，ユーザの評価傾向を学習し，インタラクションに反映する学習モデルの提案を行った．シミュレータ上に擬人化エージェント，ボールを実装し，ボールを用いたインタラクションを行える環境を構築した．構築した環境を基に感性評価実験を行い，多様なインタラクションが創発できること，提案システムが最も良い印象を得たことを確認した．今後はロボットの選択できる行動の種類を拡大したより複雑なインタラクションを扱えるシステム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．

謝辞

本研究は，一部，文部科学省科学研究費補助金（課題番号 25280100，および，25540146）の助成により行われた．

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)

[Osada 06] Osada, J., Ohnaka, S., and Sato, M.: The scenario and design process of childcare robot, PaPeRo, in *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*, ACE '06, ACM (2006)

[佐久間 14] 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)