

ユーザの嗜好に基づき
インタラクションを創発するエージェント

A Study of Interactive Agents
That Reflect User's Preference

2016

佐久間 拓人

目次

第1章 序論	1
1.1 はじめに	1
1.2 インタラクション	1
1.2.1 HCI	1
1.2.2 HRI	2
1.2.3 HAI	2
1.3 本稿におけるインタラクション	3
1.4 本研究の目的	4
1.5 本稿の構成	5
第2章 インタラクションの嗜好パターン学習および創発手法	7
2.1 はじめに	7
2.2 インタラクションモデル	7
2.3 All-Combinatorial N-gram	11
2.3.1 インタラクションルールの更新	13
2.3.2 エージェントの出力	15
2.4 「短音」の記号を用いた実験	17
2.4.1 感性評価	19
2.4.2 感性評価の結果について	20
2.4.3 創発されたインタラクション	22
2.4.4 ユーザ報酬付与傾向の検証	24

2.4.5	記号出力傾向の検証	27
2.4.6	インタラクション長の違いにおける有効性の検証	28
2.4.7	確率的な出力	31
2.4.8	確率密度の積算	32
2.5	「絵」の記号を用いた実験	34
2.5.1	感性評価	36
2.5.2	創発されたインタラクション	38
2.5.3	ユーザ報酬付与傾向の検証	43
2.5.4	記号毎の出力傾向	46
2.6	おわりに	47
第3章	ACN およびインタラクションモデルの拡張	49
3.1	はじめに	49
3.2	本章におけるインタラクション	51
3.3	ACN の改良	53
3.3.1	インタラクションルールの更新	55
3.3.2	目標状況の決定	56
3.4	感性評価実験	58
3.4.1	計算機実装	58
3.4.2	感性評価	61
3.4.3	エージェントごとの報酬付与頻度の違い	64
3.4.4	創発されたインタラクション	65
3.4.5	まとめ	72
3.5	実験参加者間の報酬付与頻度の違い	73
3.5.1	感性評価の違い	73
3.5.2	インタラクションの違い	76

3.5.3	まとめ	76
3.6	ランダムに振舞うエージェントに対してなぜ好印象を抱くのか	77
3.6.1	実験設定	77
3.6.2	印象評価	78
3.6.3	タスク成功回数と印象	80
3.6.4	報酬付与頻度	81
3.6.5	ポジティブな印象を抱いた要因	83
3.6.6	実験参加者 C の報酬付与の推移と Scaffolding	85
3.6.7	まとめ	94
3.7	おわりに	95
第 4 章	結論	97
4.1	本稿のまとめ	97
4.2	今後の課題	98
	謝辞	99
	参考文献	100
	付録 A 発表論文目録	105
	付録 B 受賞目録	107

目 次

1.1	人同士のインタラクション	4
1.2	人-エージェント間のインタラクション	4
2.1	インタラクションモデル	8
2.2	インタラクションの流れ	10
2.3	ACN による対話系列分割例 ($T_{max} = 4$ の場合)	12
2.4	インタラクションルール R	13
2.5	実験に用いた GUI	17
2.6	報酬付与中の GUI の様子	17
2.7	感性評価	21
2.8	2 対話毎の繰り返し例 1	22
2.9	2 対話毎の繰り返し例 2	22
2.10	全対話同一例 1	22
2.11	全対話同一例 2	23
2.12	ハンドクラップ・ハンドクラップ・ハンドクラップ・太鼓例 1	23
2.13	ハンドクラップ・ハンドクラップ・ハンドクラップ・太鼓例 2	23
2.14	実験参加者 A が学習時に最も高い報酬を与えたインタラクション系列 D_A	24
2.15	インタラクション長の違い (提案エージェント)	28
2.16	インタラクション長の違い (全体傾向エージェント)	29
2.17	インタラクション長の違い (ミラーリングエージェント)	29
2.18	実験に用いた GUI	34

2.19	報酬付与画面	34
2.20	感性評価	37
2.21	単一色例	38
2.22	数珠のような並び例 1	39
2.23	数珠のような並び例 2	40
2.24	数珠のような並び例 3	40
2.25	3色をバランスよく配置した例 1	41
2.26	3色をバランスよく配置した例 2	42
2.27	3色をバランスよく配置した例 3	42
3.1	エージェントの外観	50
3.2	エージェントの行動決定プロセス	51
3.3	ACN による状況系列分割例 (4 ステップ目までの場合)	53
3.4	α_k 以前に状況が存在する D_{ij} の場合の $R_{\{\alpha_k\}}$ 選出例	56
3.5	シミュレータ上のキャラクターとボールの外観	58
3.6	手の位置	59
3.7	手を移動できる場所	59
3.8	報酬付与中の様子	59
3.9	感性評価実験の結果	62
3.10	「ボールの受け渡し」の様子	65
3.11	「ボールの押し合い」の様子	66
3.12	「バレーボールのようなやりとり」の様子	67
3.13	「ボールの引っ張り合い」の様子	68
3.14	「ボールをぶつける」様子	69
3.15	「ボールを防ぐ」様子	69
3.16	「エージェントの一人遊び」の様子	70

3.17 「エージェントのフェイント」の様子	71
3.18 積極群，消極群の感性評価実験の結果	74
3.19 エージェントLにおける積極群，消極群の感性評価実験の結果	74
3.20 感性評価実験の結果	78
3.21 本実験における Scaffolding 例	84
3.22 0分～10分における実験参加者Cの報酬付与および足場タスク発生時間	85
3.23 10分～20分における実験参加者Cの報酬付与および足場タスク発生時間	85
3.24 20分～30分における実験参加者Cの報酬付与，足場タスクおよび「ボールの受け渡し」発生時間	85
3.25 図3.22中の163秒経過時において観測された Scaffolding	87
3.26 図3.22中の225秒経過時において観測された Scaffolding	87
3.27 図3.23中の980秒経過時において観測された Scaffolding	89
3.28 図3.23中の1081秒～1086秒において観測されたインタラクション	90
3.29 図3.24中の1422秒経過時において観測された「ボールの受け渡し」	91
3.30 図3.24中の1466秒経過時において観測された「ボールの受け渡し」	92

表 目 次

2.1	実験参加者 A のインタラクション系列	24
2.2	実験参加者 A の感性評価結果	24
2.3	提案エージェントのインタラクションルール集合 R (表 2.1 の系列に関連する部分)	25
2.4	全体傾向エージェントのインタラクションルール集合 R (表 2.1 の系列に関連する部分)	25
2.5	ピアソン相関係数	26
2.6	記号毎の出力回数 (実験参加者)	27
2.7	記号毎の出力回数 (提案エージェント)	27
2.8	実験参加者とエージェント D とのインタラクション	31
2.9	実験参加者とエージェント N とのインタラクション	33
2.10	実験参加者 B のインタラクション系列	43
2.11	実験参加者 B の感性評価結果	43
2.12	提案エージェントのインタラクションルール集合 R (表 2.10 の系列に関連する部分)	43
2.13	全体傾向エージェントのインタラクションルール集合 R (表 2.10 の系列に関連する部分)	44
2.14	ピアソン相関係数	45
2.15	記号毎の出力回数 (実験参加者)	46
2.16	記号毎の出力回数 (提案エージェント)	46
3.1	最も印象が良かった・悪かったエージェント	63

3.2 エージェントごとの報酬付与頻度	64
3.3 実験参加者の報酬付与頻度	73
3.4 自由記述における内容別回答人数（実験参加者数 10）	78
3.5 ランダムエージェントにおけるタスク成功回数と自由記述（実験参加者 数 10）	80
3.6 提案エージェントにおけるタスク成功回数と自由記述（実験参加者数 10）	80
3.7 ランダムエージェントにおける報酬付与回数と頻度（実験参加者数 10） .	81
3.8 提案エージェントにおける報酬付与回数と頻度（実験参加者数 10）	81

第1章

序論

1.1 はじめに

ロボットエンターテインメント [1] が提唱されて以来、AIBO[2] を始めとした様々なペットロボットや遊ぶためのロボットが研究・開発されてきた。最近では人型ロボットも増え、QRIO[3] や Robovie[4] , Pepper[5, 6] など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 [7, 8, 9]。これらの研究は主にヒューマンロボットインタラクション (HRI: Human-Robot Interaction) に分類されてきたが、近年、インタラクション相手を機械の身体を持つロボットに限定せず、コンピュータ上の存在、あるいは家具家電にも対象を拡大した分野であるヒューマンエージェントインタラクション (HAI: Human-Agent Interaction) が活発に研究されている [10]。最近では観光地などの案内役 [11] としてのエージェント開発、家電 [12] や自動車への搭載 [13] に向けたエージェント開発が多く見られる。本研究も HAI を研究対象とし、より良いエージェント開発のための一助となることを究極の目標とするものである。

1.2 インタラクション

ここで HAI と従来のヒューマンコンピュータインタラクション (HCI: Human-Computer Interaction) および HRI との違いについて整理したい。

1.2.1 HCI

まず、HCI は人間が使用するための対話型コンピュータシステムのデザイン、評価、実装に関する研究を含む学問分野であると言われている。HCI が積極的に研究され始めた時期は Apple が Macintosh を発売した時期と近く、専門家だけではなくより多くの人々がコンピュータに触れ始めたことが契機であったと考える。そのため、どのようにすればコ

ンピュータを人間にとってより良いものとするか、がHCIの大きな焦点である。一方でHAIはエージェントの設計において人間とエージェントがどのように関わっていくか、どのようにより良いインタラクションを構築できるか、といった点が焦点であると考えられる。これはコンピュータなどのシステムを擬人化する人間の特性 (media equation[14]) をHAIでは積極的に取り扱っており、情報学や工学分野のみならず、人間同士の関係を扱う社会学や社会心理学的知見がエージェント設計に用いられることから示唆される。

1.2.2 HRI

HRIはロボットそのものを開発している研究も多く見られる。有名なロボットとしては本田技研工業のASIMO[15]やNECのPapero[16]などがある。家庭用ロボットとしては、セラピーロボット「パロ」[17]、支援ロボット「wakamaru」[18]など使用用途・目的に特化したロボットの研究・開発が盛んに行われている。また、宇宙空間[19]・海中[20]や災害現場でのレスキュー用[21]といった極限環境作業用ロボットや医療用[22]といった人々の目に触れにくい領域で活動するロボットも研究・開発されている。

一般にHRIでは物理的的身体をもつロボットのみをインタラクションの対象としており、扱うインタラクションもまた身体性に注目している研究が多く見られる[23, 24, 25, 26]。一方、HAIにおける「エージェント」は山田[10]が「人間という外界とインタラクションを持つ自律システム」と述べているように、インタラクション相手は物理的的身体を持つモノに限定されず、仮想的な存在やシンボリックな表象、さらに人間もエージェントとみなしインタラクションの対象としている。そのため、HRIはHAIに内包された分野であると考えられる。また、HAIの特徴的な視点として、なぜロボットでなければいけないのか(なぜ物理的的身体が必要なのか)といった問いがあるが、HRIにおいてはそもそも提起すらされないであろう。

1.2.3 HAI

HAIが扱うインタラクションには大きく分けて三種類ある。一つは「人間とロボットのインタラクション」であり、これは従来のHRIと一致する分類である。次に「人間と仮想エージェントのインタラクション」である。これまで、ロボットのように実体を持たない仮想的なエージェントとして、ソフトウェアエージェント、インタフェースエージェント、ビリーバブルエージェント[27]、ECA (Embodied Conversational Agents) [28, 29]、擬人化エージェント[30, 31]など様々な名称が用いられてきた。本稿ではソフトウェアエージェント、ECA、擬人化エージェントをまとめたものを「仮想エージェント」と呼称する。人間と仮想エージェントのインタラクションを対象とした研究がHAIでは活発に行われており[32, 33, 34, 35]、本研究も人間と仮想エージェントとのインタラクションを主として取り扱う。最後に「エージェントを介した人間と人間のインタラクション」であり、エージェントメデイエイテッドコミュニケーション (AMC: Agent-Mediated Communication) [36]とも呼ばれる。AMCは単純に仮想エージェントを通じて多数の

人々がコミュニケーションを行う形式 [37] から、未知の存在に対して人格性（らしさ）を帰属させる過程を、未知の存在をエージェントとして仮定していると捉え、AMC とする研究も見られる [38]。すなわち、物理的な実体をもったロボット、あるいはCG などによって視覚的な実体をもった仮想エージェントといったように存在が客観的事実として定義されるもの以外にも、人自身による認識に基づいて定義される存在も「エージェント」として HAI は許容すると言える。

多種多様な仮想エージェントの中でも、本研究は特に遊び相手としてのエージェントに着目している。遊びを目的としたエージェントには何が求められており、どう自律的に行動すべきだろうか。

馬場ら [39] は人とロボットが相互に「遊び合う」関係の形成に着目し、ロボットが人に遊ばれるだけでなく、ロボットも人で遊ぶようなプラットフォームを開発している。人が一方的にロボットで遊ぶのではなく、ロボットも自身の目的へのアプローチとして人を利用することで、インタラクションは作りこみの域を超え創発されるようになり、より遊びを楽しめるのではないかと指摘している。栗山ら [40] は子どもを模したキャラクターをシミュレータ上に構築し、キャラクターとのやりとり遊びを通じて、やりとりを共創するしくみを提案している。人の応答を観察し、応答予測を形成し、予測した応答を確かめるように行動することでやりとりが繰り返される。十分に応答が確認できた場合は異なる行動を取るようにすることで多様なやりとりが生まれることを栗山らは確認している。これらを踏まえ、遊び相手としてのエージェントにはインタラクションの創発性が求められていると考える。加えて、人は遊び相手を自由に選べる場合、自身にとって良好なインタラクションができる相手を選ぶのではないだろうか。

本研究では遊びを目的としたエージェント開発において良い印象（好みである、楽しいなど）を感じるインタラクションの創発が重要であると考え、では「ユーザが良い印象を感じるインタラクション」とは具体的にどのようなインタラクションだろうか。音声を用いたインタラクションであれば会話内容、身体を用いたインタラクションであればエージェントの動きなどが印象に大きな影響を与えるだろう。しかしながら、エージェント設計者はユーザがどのようなインタラクションに対して好印象を抱くかについて事前に把握できない。そのため、本研究ではユーザ自身にエージェントへ報酬を与えてもらうことで、ユーザの好みをエージェントに学習させることとした。

1.3 本稿におけるインタラクション

本稿では人-エージェント間のインタラクションとして「やりとり」の後に人が「報酬」をエージェントに与えるものを対象としている。人同士のインタラクションにおいて「やりとり」の後に「評価」を互いに与え合うインタラクションがある。例えばジャムセッション後、「あそこは惜しかった」「あの音はすごく良かった」の感想を言い合い、再度ジャムセッションを行った際に先ほどの評価を反映したセッションを行うことでより良い「やりとり」となる。これによりお互いに良い印象を抱き、やりとり相手と良好



図 1.1: 人同士のインタラクション

な関係となるのではないかと考える。



図 1.2: 人-エージェント間のインタラクション

我々はこの人同士のインタラクション構造を人-エージェント間のインタラクションに適用することで、インタラクション相手であるエージェントに対する印象向上を狙えないかと考えた。すなわち、人は自身の好みをインタラクションの評価という形でエージェントに伝え、エージェントは人からの評価を「報酬」として学習に用いる。これにより、人が好むインタラクションを創発することが出来る学習エージェントとなる。なお、エージェントからも自身の好みを人に伝える形が人同士のインタラクションに近く、将来的には目指すべきであると考えが、本稿においては「報酬」は人からエージェントへ一方的に与えられるものであるとする。

なお、より具体的なインタラクション内容に関しては2章、3章にて述べる。

1.4 本研究の目的

本研究では、ユーザが与えた報酬を学習し、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が好むインタラクションを創発し、遊び相手であるエージェントに対する印象を向上させることを目的としている。ヒトの好みは非常に抽象的であり、分析においては「ある場面において好むモノ・コト」を対象とすることで具体化される。本研究では特に「やりとり遊び」に着目し、ユーザの好みを具体化、分析の対象とする。また、「やりとり遊び」の特徴として Ratner らは、登場する要素が少なく型にはまっていること、明確な繰り返し構造を持つことを挙げている [41]。本研究ではインタラクション中に出現するパターンの学習手法として All-Combinatorial N-gram (ACN) を提案する。これはパターン性のあるインタラクションに対してユーザが持つ好みのパターンを、ユーザからの報酬を基に動的に学習する手法であり、ACN を用いてインタラクション中

のパターンを抽出し，ユーザの好みを反映したインタラクショナルルールを作成する．

本稿では提案した手法およびシステムに対する評価として，感性評価実験及び提案手法の一部を変更したシステムによる複数の検証実験を行った．また，提案手法の有効性を生成されたインタラクショナルルール及び検証実験の結果から総合的に評価・検証する．

1.5 本稿の構成

以下に本稿の構成を示す．

1章：序論

本研究の背景，目的，アプローチなどを述べるとともに，本稿の構成を示す．

2章：インタラクシオンの嗜好パターン学習および創発手法

本研究において提案するパターン学習手法，All-Combinatorial N-gram (ACN) の詳細を述べる．また，本研究で取り扱うインタラクシオンにおける適用方法を述べる．ACNの有効性を確認する実験を，単純なやりとり遊びを題材としたインタラクシオン環境で行い，その結果と考察を述べる．やりとり遊びとして，両者が記号を交互に選び繋げていく遊びに着目する．本稿では最終的に出来上がった記号の系列を評価の対象とし，ユーザが好む系列を創発すること目的とする．記号は短音を用いたものと図形を用いたものの2種類用意し，それぞれにおいて実験を行った．インタラクシオン相手であるエージェントに対する印象評価，提案手法が獲得した実験参加者の報酬付与傾向の検証および創発されたインタラクシオンへの提案手法の寄与などを基に提案手法の有効性を述べる．

3章：ACN およびインタラクシオンモデルの拡張

2章において提案してACNをより一般的なインタラクシオンへ適用可能なよう拡張したものを提案する．また，2章と比べより複雑かつ自由度の高いインタラクシオン環境を基に実験を行い，拡張されたACNの有効性を検証する．3章ではボールを使ったやりとり遊びを題材としたインタラクシオン環境に着目し，ユーザとエージェント間でどのようにボールがやりとりされるか，その状況の遷移を評価の対象とし，ユーザが好むやりとりを創発することを目的とする．実験では実験参加者とエージェントが自由にやりとりを行い，一定時間後にエージェントに対して抱いた印象を調査する．印象調査の結果および両者間で行われたインタラクシオンを基に提案手法の有効性を述べる．特に，創発されたインタラクシオン，実験参加者の報酬付与頻度に着目して考察する．また，比較対象として用意したランダムに行動を決定するエージェントに対する謎の好印象の原因を追求する追実験も行った．追実験の結果を基に，本稿で取り扱っているようなユーザが教師としてエージェントに接するインタラクシオンにおいて重要な要素である Scaffolding についても考察する．

4章：結論

本研究の結論および今後の課題を述べる．

第2章

インタラクションの嗜好パターン学習および創発手法

2.1 はじめに

本章ではインタラクションにおけるパターン学習手法 All-Combinatorial N-gram (ACN) の提案および単純なやりとり遊びを題材としたインタラクション環境における有効性の検証を目標とする。やりとり遊びとして、両者が記号を交互に選び繋げていく遊びに着目し、記号は「短音」を用いたものと「絵」を用いたものの2種類用意し、それぞれにおいて実験を行った。

2.2 インタラクションモデル

本章ではなるべく単純なインタラクション環境を用意し実験に用いることとした。インタラクション自体の複雑さが調査対象に与える影響を抑えるためである。本稿におけるインタラクションモデルは以下の条件を満たす。

- ユーザとエージェントは記号を出力とし、インタラクションを行う
- ユーザはエージェントと行ったインタラクションに対し報酬を与え、エージェントはそれを動的に学習し、インタラクションルールを作成
- エージェントはインタラクションルールを基にユーザにとってより良いインタラクションを創発

図2.1に本研究におけるインタラクションモデルを示す。本研究ではユーザの出力及びエージェントの出力はインターフェースを通して両者に伝わるものとする。

このユーザとエージェントの出力の系列をインタラクション系列とする。また、本章ではインタラクションに以下の制約を設けた。

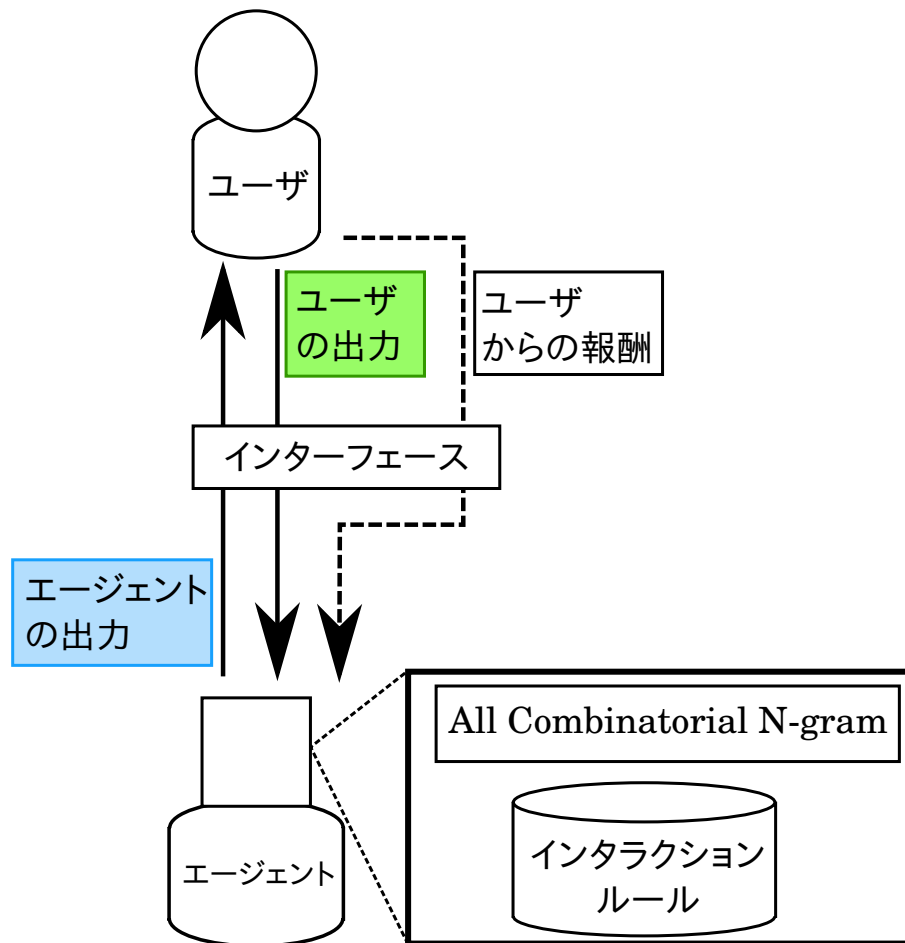


図 2.1: インタラクションモデル

- ユーザ, エージェントのどちらもあらかじめ設定された出力記号集合の中から記号を選び出力する
- ユーザとエージェントは交互に記号を出力する
- ユーザとエージェントは同じ回数だけ記号を出力する

これらの制約を満たすインタラクション例として音楽セッションにおける掛け合いが考えられる．すなわち，記号を何らかの楽器の音とし，お互いが交互に音を出すことで制約を満たす．エージェントには All-Combinatorial N-gram を実装し，ユーザ毎のインタラクションルールを保持するものとする．エージェントはユーザの出力とインタラクションルールに基づき出力を決定する．またユーザは一定回数の出力毎にエージェントとのインタラクションに対して報酬をインターフェースから与える．なお，本章ではユーザとエージェントが順番に1回ずつ出力することを1対話と呼び，規定 (T_{max}) 回対話を行うことを1インタラクションと呼ぶ． T_{max} は1インタラクションの長さであり，任意の非負整数である． T_{max} の本章における具体的な値は2.4章にて述べる．

図2.2にエージェントにおけるインタラクションの流れを示す．本章では t 対話目におけるユーザの出力を CU_t , エージェントの出力を CS_t とする．また，最大インタラクション

回数 S_{max} は任意の非負整数であり，インタラクション回数を表す変数 $S (1 \leq S \leq S_{max})$ が S_{max} と同値になった時インタラクションを終了する． S_{max} の本章における具体的な値は2.4章にて述べる．

インタラクションは大きく分けてインタラクションフェーズと報酬フェーズの2つから成る．インタラクションフェーズにおいてユーザとエージェントは記号によるインタラクションを行う．インタラクションフェーズ後の報酬フェーズにおいて，ユーザが与えた報酬を基にインタラクションルールの更新を行う．本章では，ユーザは創発された記号系列中のパターンに対して報酬を与えるとした．創発された系列にユーザが主観的に良いと思うパターンが含まれていた場合，ユーザは高い報酬値をエージェントに与える．報酬値 $U (-X \leq U \leq X)$ 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，本章における具体的な値及び報酬付与方法は2.4章にて述べる．このように，インタラクションフェーズと報酬フェーズを繰り返すことで，エージェントはユーザの報酬付与傾向を反映したインタラクションルールを獲得する．インタラクションルールは報酬フェーズの度に更新され，ユーザの報酬付与傾向を動的に獲得する． CS_t までの対話系列 $(CU_1, CS_1, \dots, CU_t)$ を，獲得したインタラクションルールに入力することで確率的にエージェントの出力 CS_t を決定する．単にユーザが高い報酬を付与したインタラクション系列を記憶・出力するだけでなく，確率を用いることで過学習を回避しつつユーザの報酬付与傾向を反映したインタラクションを創発することが期待できる．また，佐藤ら [42] は「面白さ」の演出には適度なランダム性が必要であると指摘しており，完全にランダムでも完全に適応してもユーザからの印象は下がる傾向にあると考えられる．そのため，本研究では確率的にエージェントの出力を決定できる学習モデルを提案する．

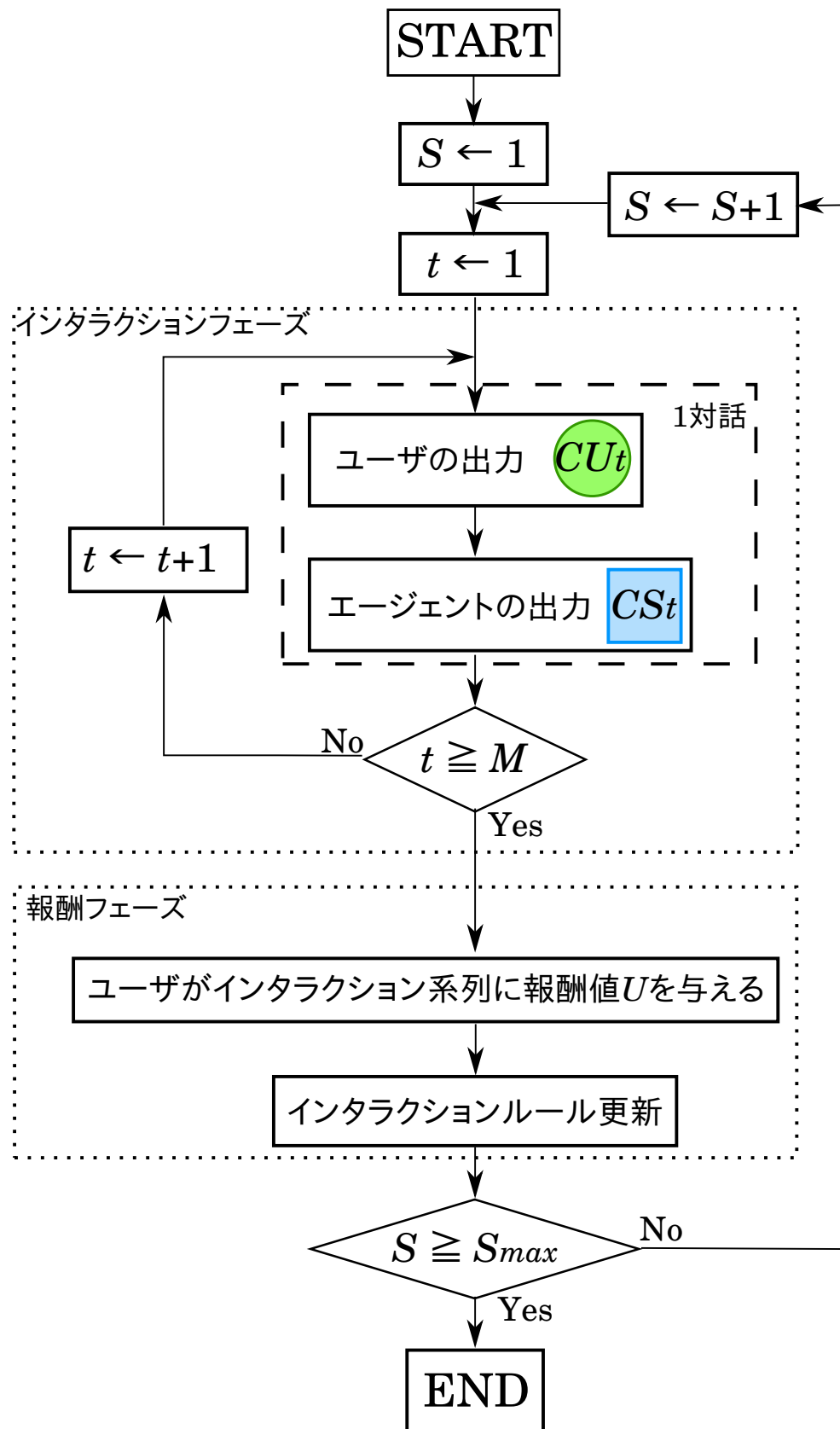


図 2.2: インタラクションの流れ

2.3 All-Combinatorial N-gram

ユーザは T_{max} 対話毎にユーザの好みに従いエージェントに報酬を与える。ユーザが与えた報酬値を1インタラクション全体への報酬としてのみ捉えたと、報酬付与時にユーザがどの対話系列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エージェントは把握出来ない。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) を使用する。ACN とは N-gram 表現に基づきインタラクション系列を分割し、学習する手法である。N-gram を用いた学習は bi-gram や tri-gram など、N の値を固定するものが多い [43, 44]。ACN では N-gram 表現によって生成可能な組み合わせを全て出力することにより、ユーザの報酬付与傾向を詳細に獲得することが出来ると考える。本章では、ACN による学習は対話単位で行うものとした。すなわち、最小単位1対話とし N の値は1から T_{max} まで変化させ、それぞれにおいて1インタラクションを分割する。そのため、対話系列は最大 $T_{max}(T_{max} + 1)/2$ 通り出力される。図 2.3(a) から 2.3(d) に $T_{max} = 4$ の場合の分割例を示す。ただし、この例では対話系列 D が重複する場合は考慮していない。

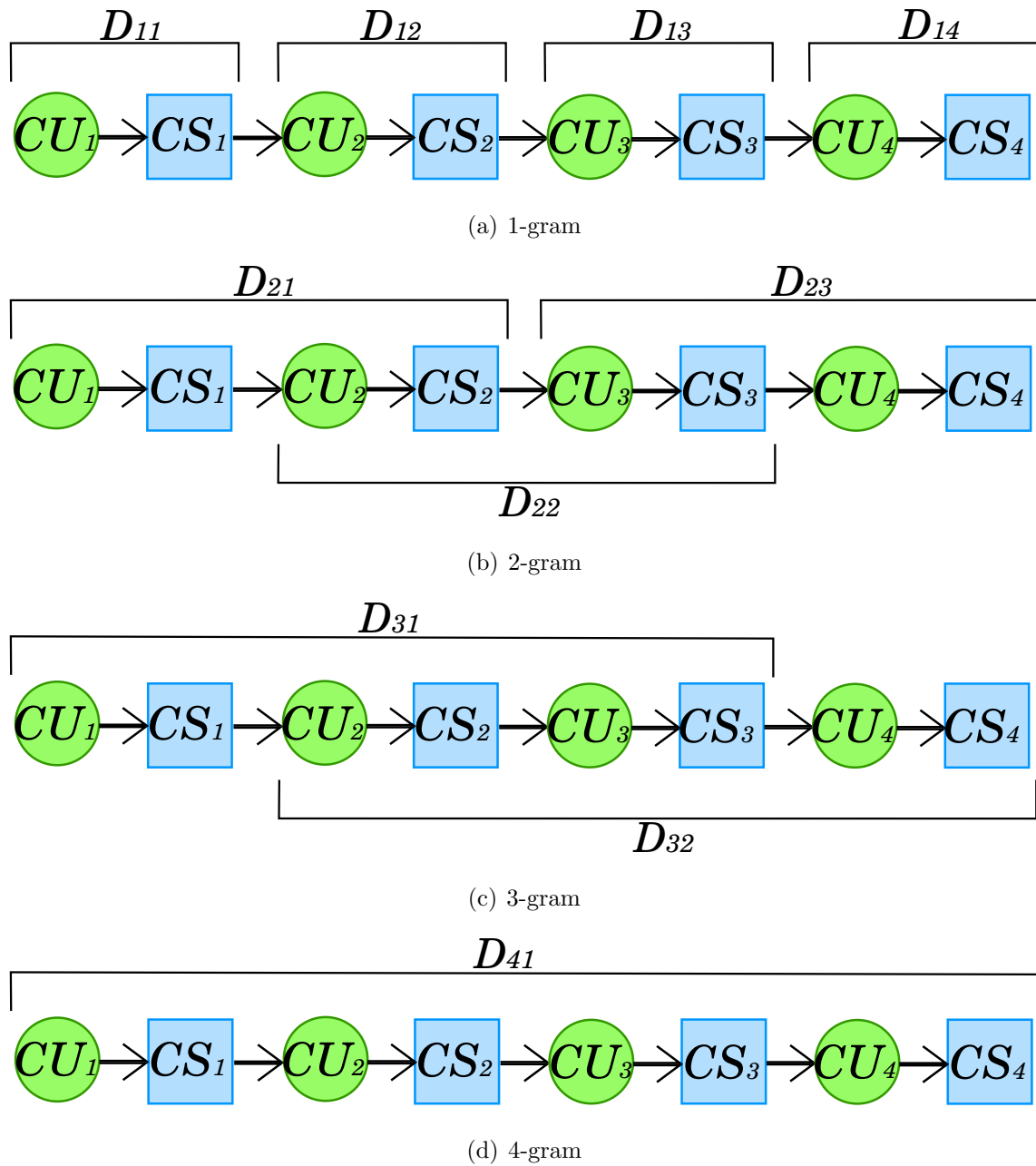


図 2.3: ACN による対話系列分割例 ($T_{max} = 4$ の場合)

2.3.1 インタラクションルールの更新

本稿では生成された N-gram モデルの全集合をインタラクションルール R と呼ぶ。 R は今までの経験を全て保持しており、報酬フェーズの度に更新される。 R の概要を図 2.4 に示す。

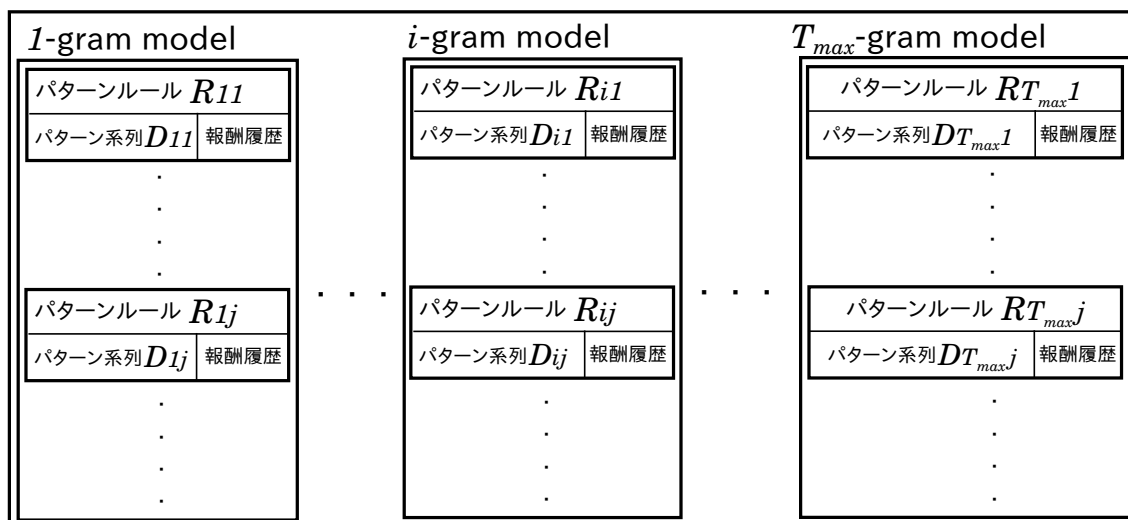


図 2.4: インタラクションルール R

また、 R の更新とは N-gram モデル内の各対話ルールの報酬値履歴を更新することとする。ここで対話ルール R_{ij} とは、ACN による分割において出力された対話系列 D_{ij} と D_{ij} の過去の報酬履歴を格納しているものとし、N-gram モデルとは対話ルールの集合である。対話系列 D_{ij} のある報酬値 P は式 (2.1) で算出される。

$$P = \frac{U \times A_{ij}}{T_{max} - \|D_{ij}\| + 1} \quad (2.1)$$

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値、 A_{ij} は D_{ij} が 1 インタラクションに出現した回数、 T_{max} は規定対話数、 $\|D_{ij}\|$ は D_{ij} の対話系列長を表す。

ユーザからの報酬をそのまま対話系列 D_{ij} に与えた場合、短い対話系列ほど何度も報酬付与されやすい特性上、対話系列の報酬値の振れ幅が大きくなり、ユーザの報酬付与傾向を正確に捉えることが出来ない。また、1 インタラクション中に何度も出現した対話系列はそれだけユーザからの報酬への影響が大きいと考えられる。そこで、式 (2.1) は対話系列の長さ ($T_{max} - \|D\|$) によって対話系列への報酬の絶対値が割引かれ、1 インタラクション中の出現回数だけ割増されるように定義されている。これにより、対話系列長の違いによる報酬値の振れ幅を減らし、かつ出現回数の多い対話系列へはより大きい報酬値 P を与えるため、正確にユーザ報酬付与傾向を捉えることが出来る考える。

なお、本インタラクションモデルにおいてユーザが報酬を与えるタイミングは一定であるため、ユーザが報酬を与えたタイミングに応じて報酬値に重み付けをする手法は採

用しなかった。インタラクションモデルを変更し、ユーザが自由なタイミングでエージェントに報酬を与えることが出来る場合においては、ユーザが報酬を与えたタイミングに近い対話系列ほど報酬値を大きくする手法が有効であると考えられる。これらの検証については3.4節にて述べる。

2.3.2 エージェントの出力

エージェントの出力 CS_t は獲得したインタラクシヨンルール R 及び、 CS_t までの対話系列 $(CU_1, CS_1, \dots, CU_t)$ によって決定される．エージェントは報酬フェーズにおけるユーザからの報酬を最大化するような出力記号を選ぶ．

ユーザとエージェントの出力記号の集合を α とし、出力 CS_t の候補を $\alpha_k \in \alpha$ とすると、エージェントは 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. 下式を満たす集合 $R_{\{\alpha_k\}}$ を求める．

$$\begin{aligned} R_{\{\alpha_k\}} = & \{ \forall R_{ij} \in \mathbf{R} \mid t - \|D_{ij}\| \geq 0 \wedge (CU_{t-\|D_{ij}\|}, CS_{t-\|D_{ij}\|}, \dots, CU_t, \alpha_k) = D_{ij} \} \\ & \cup \{ \forall R_{ij} \in \mathbf{R} \mid t - \|D_{ij}\| < 0 \wedge (CU_1, CS_1, \dots, CU_t, \alpha_k, *, \dots, *) = D_{ij} \} \end{aligned} \quad (2.2)$$

なお、 D_{ij} は対話ルール R_{ij} の持つ対話系列を、 $*$ は任意の出力記号を表す．式 (2.2) により CS_t までの対話系列と一致する部分がある対話ルールの集合が記号毎に出力される．

2. 出力記号 α_k の総報酬予測値 E_{α_k} を下式にて求める．

$$E_{\alpha_k} = \sum_{r \in R_{\{\alpha_k\}}} F(r) \quad (2.3)$$

$$F(r) = \begin{cases} \mu(r) \times \frac{1}{\sqrt{2\pi}\sigma(r)} & (\sigma(r) \neq 0) \\ \mu(r) & (\sigma(r) = 0) \end{cases} \quad (2.4)$$

ここで r は $R_{\{\alpha_k\}}$ の要素の対話ルールを表し、 $\mu(r)$ は r が持つ報酬履歴に含まれる報酬値の平均値を、 $\sigma(r)$ は標準偏差を表す．なお式 (2.4) は正規分布の確率密度関数 (式 (2.5)) を参考に、平均値の確率密度 $f(\mu(r))$ を $\mu(r)$ に積算した．

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (2.5)$$

人の感性による報酬はあいまいであり、その時の心情や状況などが影響するため同一のインタラクシヨンに常に同じ報酬が与えられるとは限らない．しかし、ある程度のばらつきはありつつも好みのインタラクシヨンには平均して高い報酬を与えると考えられる．本稿では人の報酬のばらつきは正規分布に従うと仮定した．確率密度を対話ルールの持つ報酬履歴の標準偏差から求め報酬値の平均値に積算することにより、標準偏差が大きい対話ルールは報酬値の平均値の絶対値が小さくなり、エージェントの出力決定への影響が小さくなる．相対的に標準偏差が小さい対話ルールは報酬値の平均値の絶対値が大きくなるため、エージェントの出力決定へ大きい影響を及ぼす．

以上の2ステップで算出された出力記号候補 α_k の総報酬予測値 E_{α_k} から相対的に α_k の生起確率 OP_k を算出し, OP_k に基づいて確率的に CS_t を出力する.

$$OP_k = \frac{E_{\alpha_k}}{\sum_{\alpha_k \in \alpha} E_{\alpha_k}} \times 100 \quad (2.6)$$

これにより, CS_t はユーザから高報酬が得られる確率の高い記号となる. なお, 抽出された対話ルールが一つも無い場合 ($R_{\{\alpha_k\}} = \phi$ for $\forall \alpha_k \in \alpha$) はランダムに CS_t を出力する.

2.4 「短音」の記号を用いた実験

本章で提案したエージェントの有効性を確認するため、感性評価実験を行った。感性評価実験で取り扱う記号は2種類用意し、それぞれ十分な期間を空けて実験を行った。本章では取り扱った記号の種類毎に実験を取りまとめて記述する。なお、各実験において感性評価によってユーザの好むパターンのインタラクションを創発出来るかを検証し、次に実験時のインタラクション系列の検証によってインタラクションを通じてユーザ報酬付与傾向を獲得出来るかを確認する。

本実験ではユーザとエージェントの出力記号集合 α を {シンバル, ハンドクラップ, 太鼓} の3つの短音（及びそれを表すアイコン）からなる集合とした。ユーザとエージェントのインタラクションにはGUIを用いた。図2.5に本実験で用いたGUIを示す。

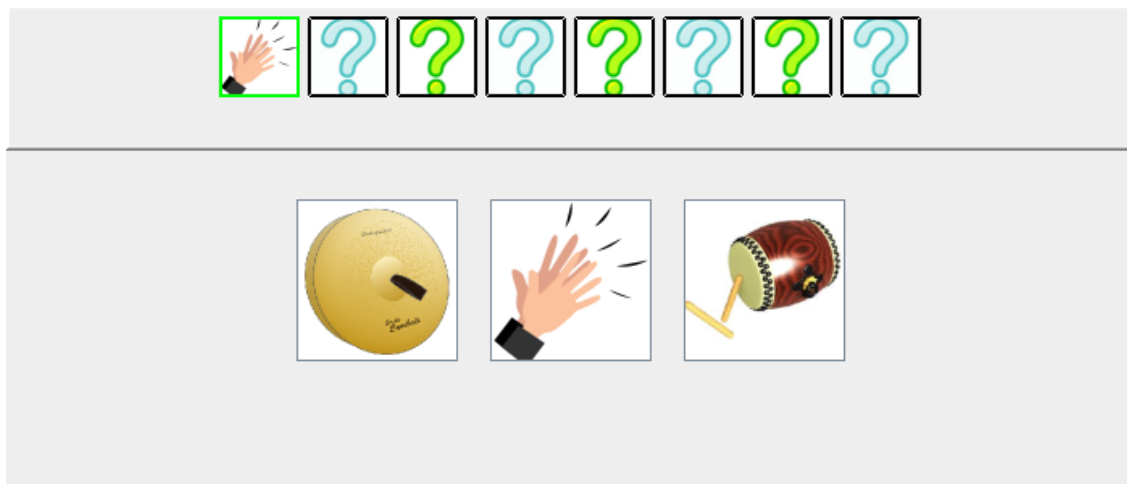


図 2.5: 実験に用いた GUI

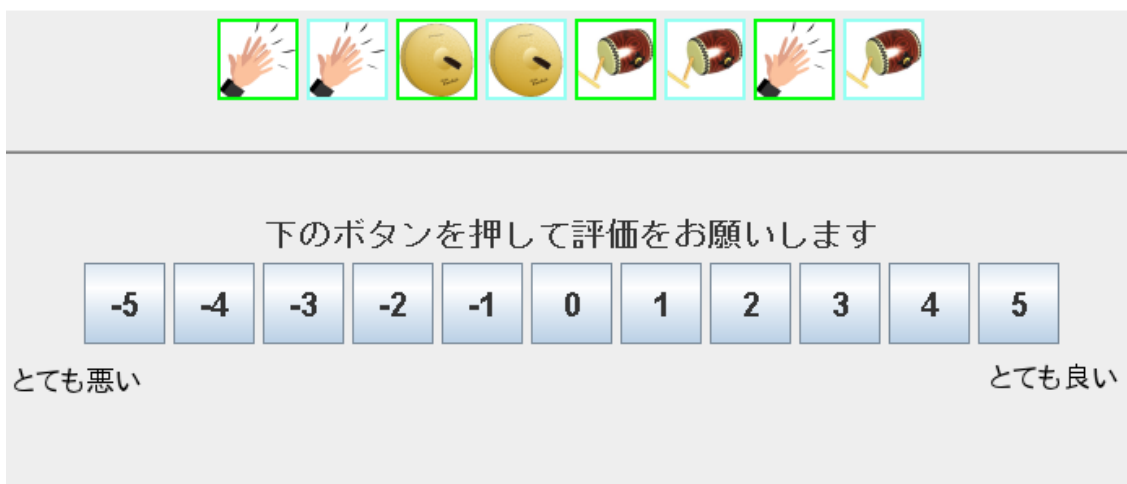


図 2.6: 報酬付与中の GUI の様子

インタラクションフェーズにおいてユーザは GUI 下段に描かれているアイコンをクリックすることで記号 (短音) を出力し, ユーザの出力から一定時間後にエージェントが出力する. ユーザ及びエージェントが出力を行うと GUI 上部に両者の出力履歴が描画される. また, 規定回数の T_{max} 対話が終了すると1インタラクションが終了し, 報酬付与画面に移行する (図 2.6). ユーザが下段に表示されている数字の描かれたボタンを押すことで, 上段に表示されている1インタラクションに対して報酬を与えたとする. ユーザからの報酬値 U は-5 (とても悪い) から+5 (とても良い) の間の11段階 ($X = 5$) とした.

なお報酬付与画面に遷移した直後に, 画面上部に描画された直前のインタラクション系列をリズム的特徴を持たせずに再生し, 実験参加者に聴取させることでインタラクションに対して報酬を与えさせた. リズム的特徴を排除した理由は, 本稿ではユーザ報酬付与傾向としてインタラクション系列内のパターンに対する主観的評価を扱うためである. そのため, インタラクションにおけるリズム的特徴であるユーザ及びエージェントの応答時間の影響を排除した.

実験参加者として20代男性13名にエージェントとインタラクションを行わせた. インタラクション回数に制限は設けず, 実験参加者は自由な回数エージェントとインタラクション出来るものとした. なお, 提案エージェントは本来動的に学習するものであるが, この実験では比較対象として動的に学習出来ないものを含むため, 実験は静的環境で行った. すなわち, あらかじめ各実験参加者に未学習の提案エージェントとやりとりさせ, 提案エージェントに各実験参加者のインタラクションルールを獲得させた. なお, 実験参加者には「良いインタラクションだと思ったら高い報酬値を, 悪いと思ったら低い報酬値を押して下さい」と指示し, インタラクションフェーズと報酬フェーズを20回繰り返す ($S_{max} = 20$) てもらった. また, 感性評価実験は学習時に受けた印象の影響を考慮し, 学習時から数日間空けた上で行った.

2.4.1 感性評価

実験参加者には4つのエージェントとやりとりをしてもらい、やりとり終了後に感性評価をもらった。なお、本節における実験では1インタラクションは4対話 ($T_{max} = 4$) とした。感性評価実験に用いたエージェントを以下に示す。

- 提案エージェント (エージェント P):
提案手法によりユーザ報酬付与傾向を学習したエージェント
- 全体傾向エージェント (エージェント W):
実験参加者全員の報酬付与傾向を学習したエージェント
- ミラーリングエージェント (エージェント M):
ユーザの出力をそのまま返すエージェント
- ランダムエージェント (エージェント R):
ユーザの出力によらずランダムに出力するエージェント

感性評価にはSD法 [45] を用い、4つの形容詞対について7段階評価で行った。評価に用いた形容詞対を以下に示す。

1. 良い-悪い(エージェントに対して)
2. 好き-嫌い(エージェントに対して)
3. 楽しかった-苦しかった(エージェントとのインタラクションに対して)
4. 成功した-失敗した(エージェントとのインタラクションに対して)

なお、これらの形容詞対はOsgoodら [45] が作成した評価性因子に分類される形容詞対から感性工学に携わる研究者と協議の未決定した。

2.4.2 感性評価の結果について

図2.7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。図2.7より、提案エージェントは感性評価の全ての項目において、他のエージェントよりも高い評価を得ていることが分かる。また、各エージェントの評価に対して t 検定による有意差検定を行った。検定の結果、提案エージェントはランダムエージェントに対して全ての項目で有意水準1%で有意にポジティブな評価がされたことがわかった（「良い」: $t = 5.096, p < 0.001$ 。「好き」: $t = 4.552, p < 0.001$ 。「楽しかった」: $t = 5.156, p < 0.001$ 。「成功した」: $t = 6.187, p < 0.001$ ）。また、全体傾向エージェントに対して有意水準1%で有意に「良い」「成功した」と評価され、有意水準5%で有意に「好き」「楽しかった」と評価された（「良い」: $t = 3.032, p = 0.006$ 。「好き」: $t = 2.504, p = 0.019$ 。「楽しかった」: $t = 2.686, p = 0.013$ 。「成功した」: $t = 2.928, p = 0.007$ ）ミラーリングエージェントに対しては有意水準5%で有意に「良い」「楽しかった」と評価された（「良い」: $t = 2.442, p = 0.022$ 。「楽しかった」: $t = 2.648, p = 0.014$ ）。このことから、提案手法による学習を行ったエージェントは他のエージェントよりもユーザにポジティブな印象を与えるインタラクションを創発することが確認された。なお、こども同士のインタラクションにおいて模倣が重要な役割を持つことを示している研究[46][47]は多いが、本実験におけるミラーリングエージェントの評価は提案エージェントと比べ有意に低い。本実験の実験参加者が全員20代男性だったことから、今回実験で扱った単純なインタラクションでは模倣遊びに対して「良い」印象や「楽しい」印象を抱くことが困難だったことが推察される。

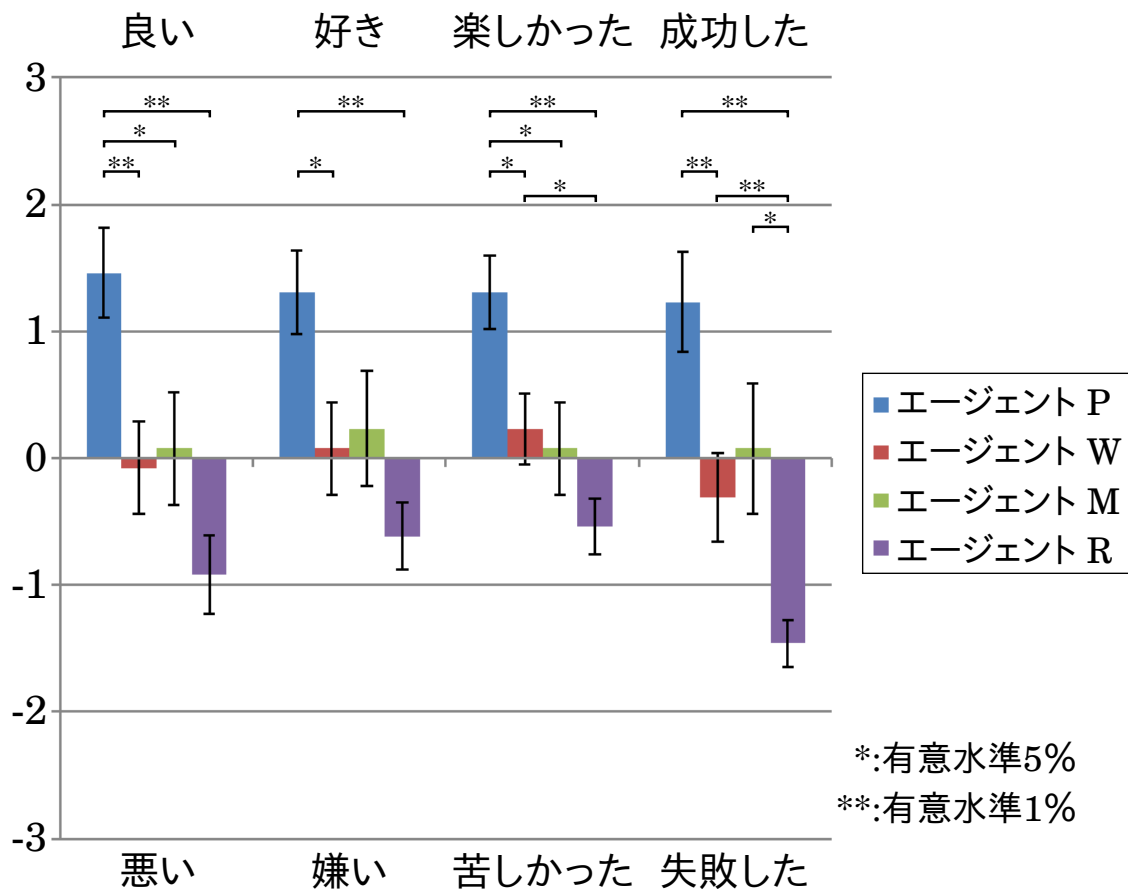


図 2.7: 感性評価

2.4.3 創発されたインタラクション

ここで、本実験環境においてどのようなインタラクションが創発されたか、抜粋して紹介する。なお、記号として用いた「シンバル」、「ハンドクラップ」、「太鼓」は便宜上それぞれ「c」、「h」、「d」と表記し、エージェントの出力は太字とする。

2 対話毎の繰り返し

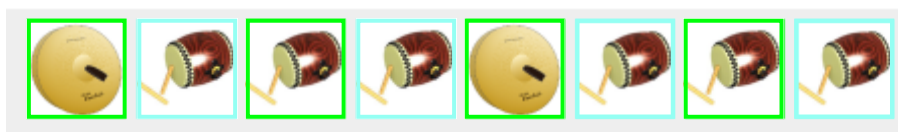


図 2.8: 2 対話毎の繰り返し例 1

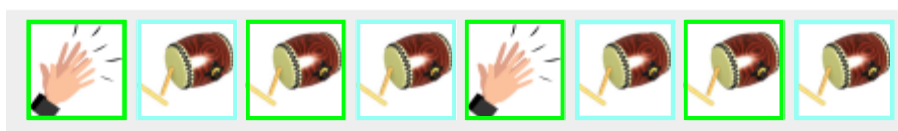


図 2.9: 2 対話毎の繰り返し例 2

図 2.8, 図 2.9 にインタラクションの例を示す。図 2.8 は (c,d,d,d,c,d,d,d) であり, 図 2.9 は (h,d,d,d,h,d,d,d) と, 2 対話毎に同じ記号が出力されていることがわかる。これを「2 対話毎の繰り返し」と呼ぶ。「2 対話毎の繰り返し」は音楽における 1 小節を意識したインタラクションであると考えられる。本実験ではエージェントの記号出力は実験参加者の出力後一定時間の間を空けて行われる。また, 最終的に 4 対話終了後にインタラクション系列を最初から一定のテンポで再生するため, 実験参加者は自然と一定のテンポに引き込まれると考える。そのため, 2 対話が 1 小節であると感じ, 1 小節毎のインタラクションを意識して報酬を与えたのではないだろうか。

「2 対話毎の繰り返し」インタラクションは多くの実験参加者で創発が見られた。特に図 2.9 のようなハンドクラップと太鼓を組み合わせたものが多く, 実験参加者から高い報酬を与えられていた。また, インタラクション系列中の全対話が同一なインタラクションの創発も多数見られた。図 2.10, 図 2.11 にインタラクションの例を示す。

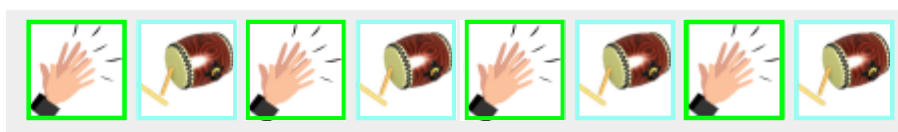


図 2.10: 全対話同一例 1



図 2.11: 全対話同一例 2

図 2.10 は (h,d,h,d,h,d,h,d) であり，図 2.11 は (h,h,h,h,h,h,h,h) である．図 2.10，図 2.11 は提案手法における ACN の性質上非常に学習しやすい系列であり，一度高い報酬を付与されるとエージェントの出力が偏りやすくなる．実験参加者によっては，はじめは高い報酬を付与したもののエージェントが極端に同一の出力しかしなくなったために低い報酬を付与し続けた様子も観測した．図 2.11 におけるインタラクション系列は特にミラーリングエージェントとのインタラクションにおいてもよく観測されており，あまり高い報酬は付与されない系列であった．

ハンドクラップ・ハンドクラップ・ハンドクラップ・太鼓



図 2.12: ハンドクラップ・ハンドクラップ・ハンドクラップ・太鼓例 1

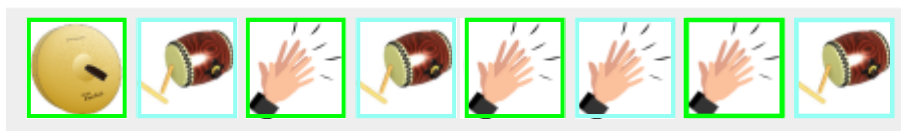


図 2.13: ハンドクラップ・ハンドクラップ・ハンドクラップ・太鼓例 2

図 2.12，図 2.13 に図 2.8，図 2.9 とは異なるタイプのインタラクション例を示す．図 2.12，図 2.13 はそれぞれ (c,d,h,d,h,h,h,d) と (h,d,h,h,h,h,h,d) となっており，小節を意識して 2 対話毎に区切ると 2 小節目がどちらも (h,h,h,d) であることがわかる．なお，どちらもある一人の実験参加者とのインタラクション中に創発したものである！「2 対話毎の繰り返し」同様，実験参加者が小節を意識したために学習されたインタラクション系列であると考えられる．1 小節目が異なるにも関わらず 2 小節目は同じ系列となっており，どちらのインタラクションに対しても高い報酬が与えられていたことから，実験参加者の好みを捉えた創発が出来ていると考える．

2.4.4 ユーザ報酬付与傾向の検証

感性評価実験時の自由記述の回答において、提案エージェントに対して「自分のパターンを反映しているように感じた」「自分の好みのパターンがよく出てきた」と記述した実験参加者がみられた。そこで、提案エージェントがユーザの報酬付与傾向を反映出来たかどうか、インタラクションの履歴を用いて考察する。典型例として図 2.7 の評価平均に最も近い感性評価を与えた実験参加者 A を取り上げる。

表 2.1 に実験参加者 A が学習時に最高報酬を与えたインタラクション系列 D_A 、及び D_A のユーザの出力を提案エージェントと全体傾向エージェントに与えた場合の創発確率が最も高いインタラクション系列をそれぞれ示す。図 2.14 に実験参加者 A が学習時に最高報酬を与えたインタラクション系列 D_A の GUI 上の表示を示す。表 2.2 に実験参加者 A の感性評価結果を示す。また、表 2.3 と表 2.4 に提案エージェントと全体傾向エージェントがそれぞれ獲得したインタラクションルール集合のうち、表 2.1 の系列に関連する部分を示す。なお、表中の「h」及び「d」はそれぞれ出力記号の「ハンドクラップ」と「太鼓」であり、太文字はエージェントの出力を意味する。また、 $F(R)$ は式 (2.4) と同様に対話ルール R が持つ報酬履歴に含まれる報酬値の平均値に尤度を積算した値であるが、報酬履歴が存在しない対話ルールの $F(R)$ は ϕ と表記した。

表 2.1: 実験参加者 A のインタラクション系列

学習時に最も高い報酬を得た インタラクション系列 D_A	D_A と同一のユーザ出力を与えた場合に 創発確率が最も高いインタラクション系列	
	提案エージェント	全体傾向エージェント
d,d,h,d,h,d,d,h	d,d,h,d,h,d,d,h	d,h,h,h,h,h,d,d

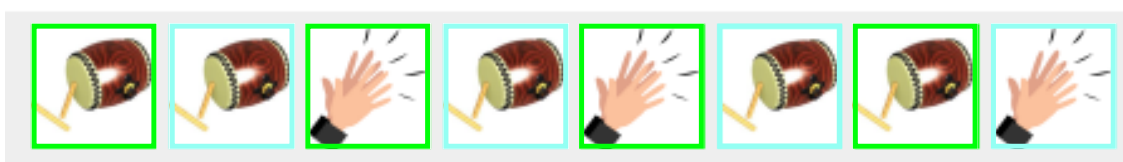


図 2.14: 実験参加者 A が学習時に最も高い報酬を与えたインタラクション系列 D_A

表 2.2: 実験参加者 A の感性評価結果

評価項目	提案 エージェント	全体傾向 エージェント	ミラーリング エージェント	ランダム エージェント
良い-悪い	1	-1	-2	-2
好き-嫌い	1	0	-2	-1
楽しかった-苦しかった	2	0	-1	-1
成功した-失敗した	2	-1	-2	-2

表 2.1 より、ユーザの出力が D_A と同一の時、提案エージェントが創発する確率が最も高いインタラクション系列は D_A と同一であることが分かる。それに比べ、全体傾向エージェントではユーザの出力が同一であるにも関わらず、創発する確率が最も高いインタ

表 2.3: 提案エージェントのインタラクシヨンルール集合 R (表 2.1 の系列に関連する部分)

インタラクシヨンルール R							
1-gram model		2-gram model		3-gram model		4-gram model	
対話系列	$F(R)$	対話系列	$F(R)$	対話系列	$F(R)$	対話系列	$F(R)$
d,d	0.2393	d,d,h,d	1.0592	d,d,h,d,h,d	1.0638	d,d,h,d,h,d,d,h	5
h,d	0.3055	h,d,h,d	1.5957	h,d,h,d,d,h	2.5	d,d,h,d,h,d,d,d	1
d,h	0.0389	h,d,d,h	1.6667	h,d,h,d,d,d	0.5	d,h,h,h,h,h,d,d	ϕ
h,h	0	h,d,d,d	0.3333	d,h,h,h,h,h	ϕ		
		d,h,h,h	-0.3333	h,h,h,h,d,d	ϕ		
		h,h,h,h	ϕ				
		h,h,d,d	-0.2564				

表 2.4: 全体傾向エージェントのインタラクシヨンルール集合 R (表 2.1 の系列に関連する部分)

インタラクシヨンルール R							
1-gram model		2-gram model		3-gram model		4-gram model	
対話系列	$F(R)$	対話系列	$F(R)$	対話系列	$F(R)$	対話系列	$F(R)$
d,d	0.2187	d,d,h,d	0.4135	d,d,h,d,h,d	1.0638	d,d,h,d,h,d,d,h	5
h,d	0.2448	h,d,h,d	0.6708	h,d,h,d,d,h	2.5	d,d,h,d,h,d,d,d	1
d,h	0.0832	h,d,d,h	0.3363	h,d,h,d,d,d	0.5	d,h,h,h,h,h,d,d	4
h,h	0.2249	h,d,d,d	0.1547	d,h,h,h,h,h	0.8865		
		d,h,h,h	0.3016	h,h,h,h,d,d	2		
		h,h,h,h	0.1850				
		h,h,d,d	0.4467				

ラクシヨン系列は D_A と異なっていることが分かる。また表 2.3, 表 2.4 の結果から, D_A (d,d,h,d,h,d,d,h) の創発確率は提案エージェントの場合約 20 % であるのに対し, 全体傾向エージェントの場合約 5 % であり, ランダムエージェントの場合約 1.2 % である。このことから, 提案エージェントは全体傾向エージェントに対して約 4 倍, ランダムエージェントに対して約 16.7 倍の確率で D_A を創発することが可能であることが分かる。以上から, 提案エージェントは実験参加者 A の傾向を反映したインタラクシヨンを創発することができていると考えられる。なお, 実験の際に実験参加者の出力が学習時と同様であった割合は約 38 % であった。このことから実験参加者らは一つの理想的な系列を事前に想定してインタラクシヨンを行っていないことが推察される。このため, 単純に学習時に最大の報酬値が得られた記号系列を最長一致で検索する方策では対応できないと考える。また, 本手法はインタラクシヨンを続けることでより良い記号系列の創発を目的としており, 動的に学習し続けることを想定しているため, その時点までの最高報酬が与えられた系列を出力する方策ではユーザからより高い報酬が与えられる系列の出力は不可能であり, 発展が望めないと考えられる。そのため, ユーザの報酬付与傾向を獲得し, ユーザの出力に応じてエージェントの出力を確率的に決定できる本手法が有効であると考えられる。

ここで, 表 2.3 と表 2.4 を比較すると, 2-gram model の $F(R)$ が大きく異なることが分かる。また, 全体傾向エージェントのインタラクシヨンルールには存在するが, 実験参加者 A のインタラクシヨンルールには報酬履歴が存在しない対話ルールがあることも分かる。インタラクシヨン系列の創発確率は式 (2.2) ~ (2.6) により算出されるため関与

する対話ルール R の $F(R)$ によって決まると言ってもよい。実験参加者 A の場合、表 2.4 の (d,d,h,d) と (h,d,h,d) と (h,d,d,h) の $F(R)$ が表 2.3 では大きく増加しており、他の対話系列とも明確に差が出ている。これらの差が実験参加者 A の報酬付与傾向として、創発されたインタラクション系列の違いを生んだと考えられる。

表 2.5: ピアソン相関係数

実験参加者	提案エージェント	全体傾向エージェント
A	0.684	-0.256
B	0.967	0.755
C	0.896	0.773
D	0.592	0.296
E	0.857	0.149
F	0.704	0.496
G	0.922	0.828
H	0.865	0.236
I	0.879	-0.679
J	0.894	0.335
K	0.920	0.577
L	0.834	0.679
M	0.824	0.437
all	0.843	0.500

ユーザの傾向はインタラクションルール中の対話ルール R から算出される $F(R)$ の傾向として表現されたと考える。学習時に実験参加者から報酬を与えられた系列に対して、その系列の創発に関与する対話ルール R の $F(R)$ を求め、実験参加者が与えた報酬値 U との相関分析を行った。分析には実験参加者 13 名全員の学習データを用いて、ピアソン相関係数を求めた。表 2.5 に実験参加者 13 名の個々のピアソン相関係数および全員のデータを使用したピアソン相関係数を示す。分析の結果、提案エージェントにおける全員のデータを使用した相関係数は 0.84 であり、ユーザ報酬値 U と $F(R)$ の高い相関が確認された。一方で、同様の分析を全体傾向エージェントに対して行ったところ相関係数は 0.50 であり、提案エージェントと比べて相関係数は明らかに低かった。各実験参加者のデータにおいてもそれぞれ相関係数を求めたところ、全実験参加者において全体傾向エージェントに比べ提案エージェントの相関係数が高いことがわかった。さらに提案エージェントにおける相関係数は 13 名中 12 名が 0.6 を超えているが、全体傾向エージェントにおける相関係数は 13 名中 4 名のみが 0.6 を超えていることがわかった。以上から、提案エージェントはユーザ報酬値と相関の高いインタラクションルールを生成することができたことが分かる。

2.4.5 記号出力傾向の検証

表 2.6: 記号毎の出力回数（実験参加者）

実験参加者の出力	合計（全実験参加者）	割合
シンバル	263	27.3%
ハンドクラップ	386	37.1%
太鼓	391	37.6%
全記号	1040	

表 2.6 に記号毎の全実験参加者合計出力回数を示す．表 2.6 より，実験参加者自身の出力は太鼓が最も多く，シンバルが最も少なかったことがわかる．シンバルの数が目立って少ないことに関しては，本実験における実験参加者の好みの偏りであると考えられる．本実験で用いた各記号の内，ハンドクラップ，太鼓と比べてシンバルは連続で鳴らす性質の楽器ではない．これはシンバルのアイコンとして用いた絵がクラッシュ・シンバル（あるいはハンドシンバル）と呼ばれる物を模したものとなっていることから，頻繁に鳴らすものではないという印象が強くなったと考えられる．そのため，表 2.6 のような出力回数の偏りが生じたと考えられる．あるいは単純に，ハンドクラップ，太鼓とくらべシンバルの音をうるさく感じたために出力が抑えられたとも考える．

表 2.7: 記号毎の出力回数（提案エージェント）

提案エージェントの出力	合計（全実験参加者）	割合
シンバル	292	28.1%
ハンドクラップ	351	33.8%
太鼓	397	38.2%
全記号	1040	

表 2.7 に事前学習時における提案エージェントの記号毎合計出力回数を示す．表 2.7 より，太鼓が最も多く，シンバルが最も少ないことがわかる．提案エージェントの出力は，実験参加者から与えられた報酬を基に決定されるため，表 2.7 のような出力の偏りは本実験参加者における記号毎の好みの偏りであると考えられる．

2.4.6 インタラクション長の違いにおける有効性の検証

前節の実験では1インタラクションを4対話 ($T_{max} = 4$) で固定していたが、1インタラクションの長さを変化させた実験を提案エージェント、全体傾向エージェント、ミラーリングエージェントを対象に行った。ランダムエージェントは2.4.2節の感性評価において全項目で評価平均値が低かった(図2.7)ため、実験の比較対象から除外した。

図2.15, 図2.16, 図2.17に $T_{max} = 2, 4, 6, 8$ における提案エージェント、全体傾向エージェント、及びミラーリングエージェントの評価結果をそれぞれ示す。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。なお、実験参加者は20代男性13名であり、実験に使用したGUI, 出力記号, 及び各パラメータは2.4節のものと同一である。

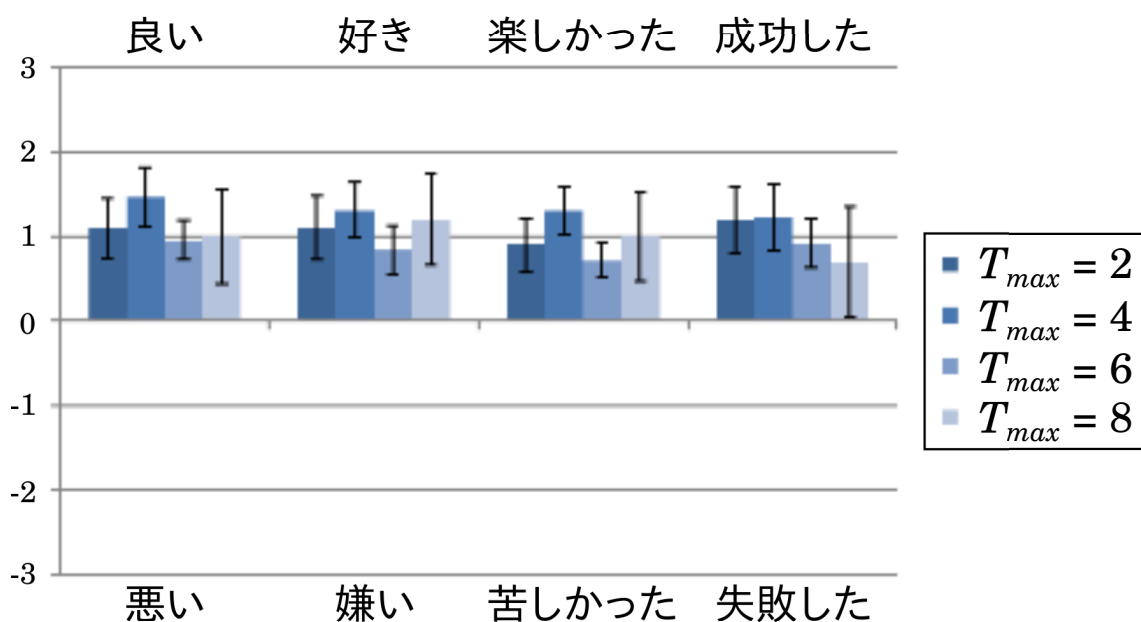


図 2.15: インタラクション長の違い (提案エージェント)

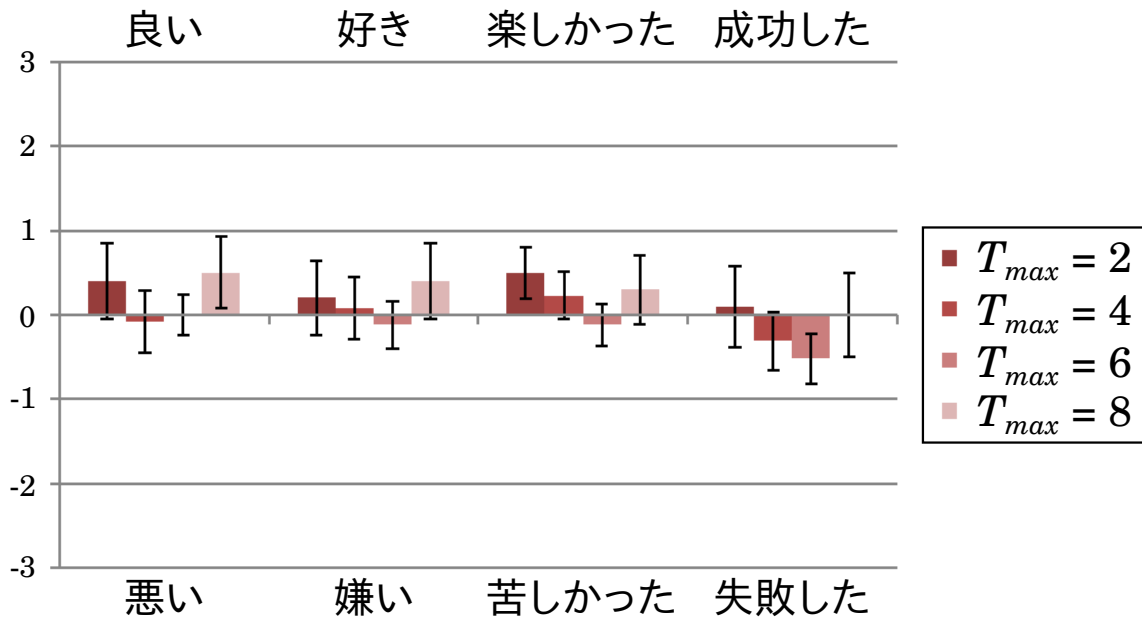


図 2.16: インタラクション長の違い (全体傾向エージェント)

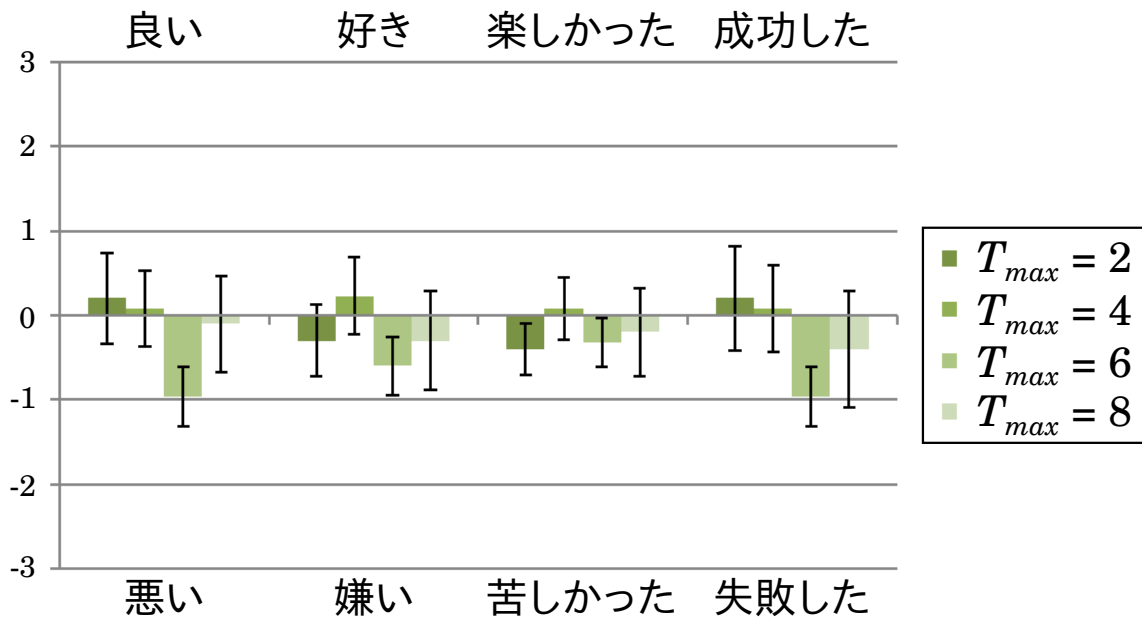


図 2.17: インタラクション長の違い (ミラーリングエージェント)

図 2.15 から図 2.17 より、提案エージェントはインタラクション長 T_{max} によらず全項目で他のエージェントより高い評価平均値を出していることが分かる。これにより、提案エージェントは様々な長さのインタラクション系列においてもユーザにとってポジティブな印象を与えるインタラクションを創発することが分かった。

図 2.15 より、「成功した」の項目において T_{max} の増加に伴い評価平均値が減少していることが分かる。実験参加者の自由記述においても $T_{max} = 8$ における学習時のインタラクションは「学習させづらい」、あるいは「あまり学習しなかった」との意見が得られた。実験参加者には学習のためにエージェントと 20 回 ($S_{max} = 20$) インタラクションさせているが、この学習のためのインタラクションが $T_{max} = 6$ や $T_{max} = 8$ の場合 20 回では少なかったと考えられる。学習回数を増やすことである程度は対処可能と考えるが、ユーザの負担を考慮すると無闇に学習回数を増やすことは好ましくない¹。

また、 $T_{max} = 6$ の実験においては実験参加者からの印象が特に低かった。この要因として、「音楽的に不自然」「1.5 小節なので中途半端で気持ち悪い」との意見が得られたことから、音楽的な要素を考慮しなかった点が考えられる。本実験は音楽の創発ではなく、あくまで「短音」のインタラクションという想定ではあったが、エージェントが出力するテンポが一定であること、報酬付与直前にインタラクションが再度再生される際のテンポも一定であったことから、自然と 1 小節の区切りを実験参加者が意識するに至ったと考える。

¹本エージェントにおけるインタラクションパターンの場合の数は $(\#\alpha)^{T_{max}}$ である

2.4.7 確率的な出力

本稿では2.2章において述べた通り，確率を用いた出力によって過学習の回避を想定している．出力を決定的に行うエージェントの場合，本手法の出現した系列にのみ報酬が与えられる特性上過学習になりやすい．このことに対する検証実験として，最も高い総報酬予測値 E_{α_k} となる記号を決定的に出力するエージェント（エージェント D）に対して学習実験を行った．実験参加者は20代の男女13名であり，エージェント D を用いてインタラクションフェーズと報酬フェーズを20回繰り返してもらった．その結果，三名の実験参加者において明確な過学習が観測された．うち一名の実験参加者は20回のインタラクション中の4回目以降17回同一の系列を生成しており（表2.8），このような場合学習時に出現した4種類の系列にのみ報酬が与えられているため，異なる系列への対応は難しいと考えられる．なお，表中の c, h, d はそれぞれ出力記号の「シンバル」，「ハンドクラップ」，「太鼓」であり．太文字はエージェントの出力を意味している．このように決定的に出力することは過学習をまねく危険性があり，確率的に出力を決定することで過学習を回避することが出来る考える．

表 2.8: 実験参加者とエージェント D とのインタラクション

S	エージェント D とのインタラクション履歴	実験参加者が与えた報酬
1	d,h,h,c,h,d,h,c	-4
2	d,c,h,h,h,h,h,h	-5
3	d,d,h,d,h,d,h,d	-5
4	d,h,h,d,h,c,h,c	2
5	d,h,h,d,h,c,h,c	3
6	d,h,h,d,h,c,h,c	1
7	d,h,h,d,h,c,h,c	2
8	d,h,h,d,h,c,h,c	-2
9	d,h,h,d,h,c,h,c	-2
10	d,h,h,d,h,c,h,c	-2
11	d,h,h,d,h,c,h,c	-2
12	d,h,h,d,h,c,h,c	-4
13	d,h,h,d,h,c,h,c	-3
14	d,h,h,d,h,c,h,c	-3
15	d,h,h,d,h,c,h,c	-3
16	d,h,h,d,h,c,h,c	-3
17	d,h,h,d,h,c,h,c	-3
18	d,h,h,d,h,c,h,c	-5
19	d,h,h,d,h,c,h,c	-5
20	d,h,h,d,h,c,h,c	-5

2.4.8 確率密度の積算

2.3.2章の式(2.4)において本手法では正規分布の確率密度を用いている。確率密度を使用しない場合、報酬値のばらつきによる過度の影響が考えられる。この検証実験として、提案手法の式(2.4)において確率密度を積算しない式を使用するエージェント(エージェントN)の学習実験を行った。実験参加者は20代の男女13名であり、エージェントNを用いてインタラクションフェーズと報酬フェーズを20回繰り返してもらった。その結果、四名の実験参加者において報酬値の影響を過度に受け学習が安定しない様子が見られた。うち一名の実験参加者とエージェントNとのインタラクションの結果を表2.9に示す。本研究では「実験参加者は好みの系列には平均して高い報酬を与え、ばらつきは正規分布に従う」と仮定しており、本来同一の系列には同一の報酬が与えられることが望ましい。しかし、表2.9における1回目、2回目のインタラクションに対する報酬は大きく違うにも関わらず前半の系列(c,d,d,c)は同一であるため(c,d,d,c)への報酬値はばらついたものとなる。同様に6回目、7回目のインタラクションでは(d,h,c,d,d,h)の系列が共通しているにもかかわらず実験参加者の与えた報酬値が大きく違うため(d,h,c,d,d,h)に対する報酬値はばらついてしまう。このような場合、共通している系列はユーザにとってどの程度の価値があるか判断できないため、報酬値の標準偏差を考慮せず平均値のみを用いることでばらつきの大きい報酬値の影響が過度に考慮された出力になる危険性が考えられる。

表 2.9: 実験参加者とエージェント N とのインタラクション

<i>S</i>	エージェント N とのインタラクション履歴	実験参加者が与えた報酬
1	c,d,d,c,c,h,h,c	2
2	c,d,d,c,h,c,d,c	-4
3	c,h,h,c,d,h,h,d	-2
4	h,h,d,d,c,c,h,h	-5
5	c,h,d,h,h,c,c,d	-1
6	d,h,c,d,d,h,c,d	2
7	d,h,c,d,d,h,h,c	-3
8	c,c,d,c,c,d,h,d	-3
9	c,h,d,d,h,c,c,h	-1
10	c,d,d,h,h,d,d,h	-1
11	h,d,d,h,c,d,d,d	-2
12	c,d,d,d,h,c,c,d	-5
13	c,c,d,h,c,h,h,c	-2
14	c,h,d,c,c,c,h,c	-4
15	d,c,h,h,d,h,c,c	-3
16	h,d,d,d,c,h,h,c	-2
17	d,h,h,c,c,h,c,c	-4
18	h,c,d,d,c,d,h,c	-3
19	c,c,h,d,c,h,d,h	2
20	c,c,d,h,h,d,d,c	-2

2.5 「絵」の記号を用いた実験

本実験ではユーザとエージェントの出力記号集合 α をビーズを模した3色の丸い図形からなる集合とした。2.4章の実験では「短音」を記号として用いたためにテンポを考慮しないことの影響が懸念された。

本実験で想定されるインタラクションのパターンとは図形の並び順からなる「絵」である。そのため、インタラクション系列内のパターンをより分かりやすく、明確にすることを目標に GUI を設計した。図 2.18 に本実験で用いた GUI を示す。

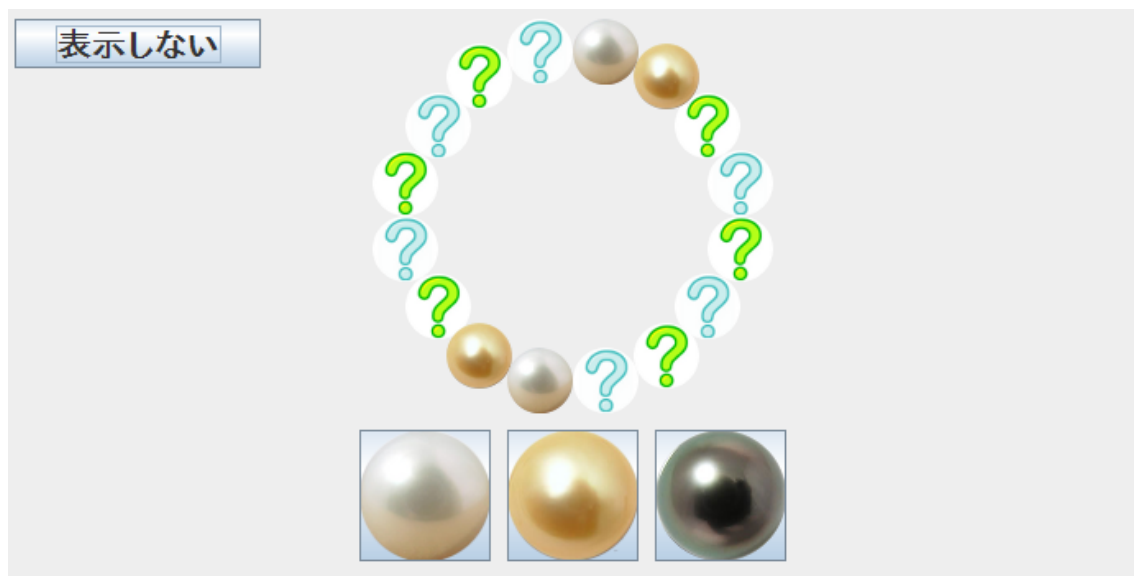


図 2.18: 実験に用いた GUI



図 2.19: 報酬付与画面

インタラクションフェーズにおいてユーザは GUI 下段に描かれているアイコンをクリックすることで記号（絵）を出力し，ユーザの出力から一定時間後にエージェントが出力する．ユーザおよびエージェントが出力を行うと GUI 上部に両者の出力履歴が円形に描画される．なお 1 インタラクションは 4 対話 ($T_{max} = 4$) とし，円形の左半分は右半分の上下左右反転コピーが描画される．円形であることで直線に記号が並んだ状態とは異なるパターンの創発が期待出来る．これは 2.4 章の実験に参加した実験参加者が，本実験においても同一のパターンを創発することによる主観評価への影響を考慮している．なお，2.4.6 節における結果から，1 インタラクションの長さは 4 対話 ($T_{max} = 4$) が最適であると判断した．また，規定回数の T_{max} 対話が終了すると 1 インタラクションが終了し，報酬付与画面に移行する．表示される報酬付与画面を図 2.19 に示す．ユーザからの報酬値 U は -5（とても悪い）から +5（とても良い）の間の 11 段階 ($X = 5$) とした．

実験参加者として 20 代男性 13 名にエージェントとインタラクションを行わせた．インタラクション回数に制限は設けず，実験参加者は自由な回数エージェントとインタラクション出来るものとした．なお，提案エージェントは本来動的に学習するものであるが，この実験では比較対象として動的に学習出来ないものを含むため，実験は静的環境で行った．すなわち，あらかじめ各実験参加者に未学習の提案エージェントとやりとりさせ，提案エージェントに各実験参加者のインタラクションルールを獲得させた．なお，実験参加者には「良いインタラクションだと思ったら高い報酬値を，悪いと思ったら低い報酬値を押して下さい」と指示し，インタラクションフェーズと報酬フェーズを 20 回繰り返し ($S_{max} = 20$) てもらった．また，感性評価実験は学習時に受けた印象の影響を考慮し，学習時から数日間空けた上で行った．

2.5.1 感性評価

実験参加者には4つのエージェントとやりとりさせ、やりとり終了後に感性評価をしてもらった。評価実験に用いたエージェントを以下に示す。

- 提案エージェント (エージェント P):
提案手法によりユーザ報酬付与傾向を学習したエージェント
- 全体傾向エージェント (エージェント W):
実験参加者全員の報酬付与傾向を学習したエージェント
- ミラーリングエージェント (エージェント M):
ユーザの出力をそのまま返すエージェント
- ランダムエージェント (エージェント R):
ユーザの出力によらずランダムに出力するエージェント

感性評価にはSD法 [45] を用い、5つの形容詞対について7段階評価で行った。評価に用いた形容詞対を以下に示す。

1. 良い-悪い(エージェントに対して)
2. 親しみやすい-親しみにくい(エージェントに対して)
3. 面白かった-つまらなかった(エージェントに対して)
4. 自由だった-不自由だった(ビーズパターンに対して)
5. 思い通りだった-思い通りでなかった
(ビーズパターンに対して)

なお、これらの形容詞対は2.4節の実験において使用した形容詞対、およびエージェントの評価に多く用いられる形容詞対から、感性工学に携わる研究者と協議の未決定した。

図2.20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。図2.20より、提案エージェントは感性評価の全ての項目において、他のエージェントよりも高い評価を得ていることが分かる。また、各エージェントの評価に対して t 検定による有意差検定を行った。検定の結果、提案エージェントは全体傾向エージェントに対して有意水準1%で有意に「良い」「親しみやすい」「思い通りだった」と評価された(「良い」: $t = 5.564, p < 0.001$ 。「親しみやすい」: $t = 4.591, p < 0.001$ 。「思い通りだった」: $t = 5.062, p < 0.001$)。ミラーリングエージェントに対しては有意水準1%で有意に「良い」「面白かった」「自由だった」と評価され、有意水準5%で有意に「親しみやすい」と評価された(「良い」: $t = 4.010, p = 0.001$ 。「親しみやすい」: $t = 2.343, p = 0.029$ 。「面白かった」: $t = 5.343, p < 0.001$ 。「自由だった」: $t = 4.667, p < 0.001$)。ランダムエージェントに対しては有意水準1%で有意に「良い」「親しみやすい」「思い通りだった」と評価され、有意水準5%で有意に「自由だった」と評価された(「良い」: $t = 3.722, p = 0.001$ 。「親しみやすい」: $t = 2.778, p = 0.010$ 。「自由

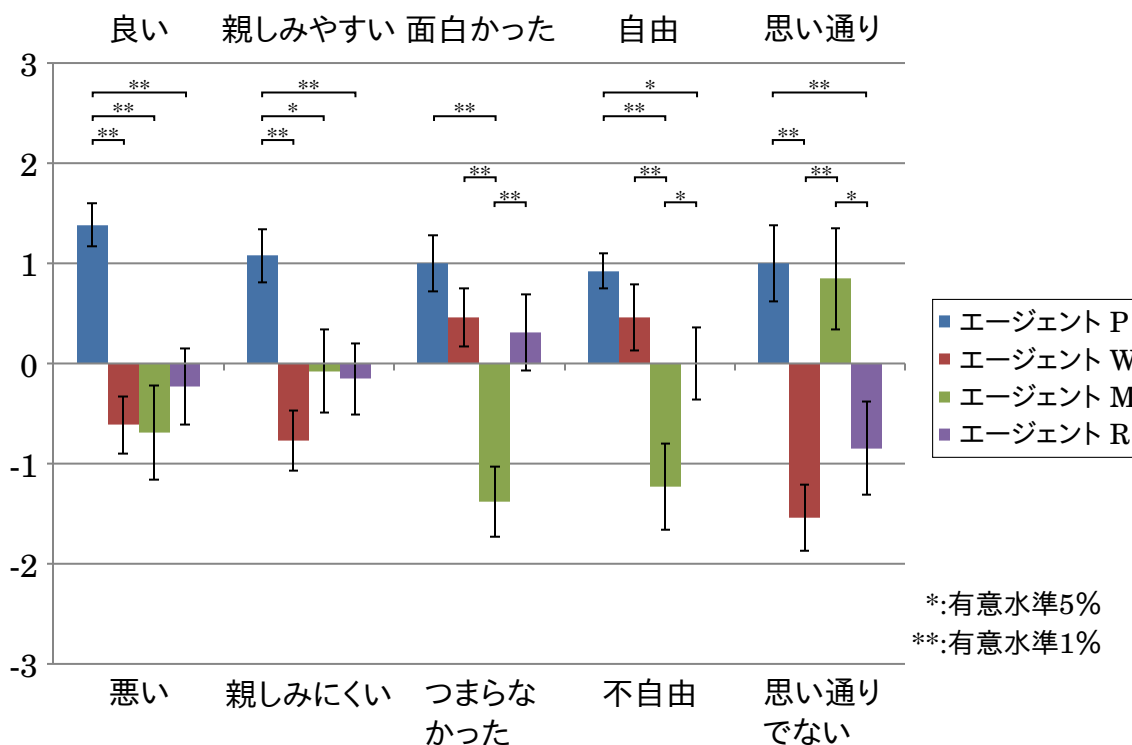


図 2.20: 感性評価

だった」: $t = 2.309, p = 0.033$ 。「思い通りだった」: $t = 3.090, p = 0.005$)。これにより、提案エージェントはユーザの与えた「良い-悪い」の報酬付与傾向を学習し、インタラクションに反映することが出来ていると考えられる。また、比較対象の全エージェントに対して有意に「親しみやすい」と評価されたことから、提案手法による学習を行ったエージェントは他のエージェントよりもユーザにポジティブな印象を与えることが確認された。なお、ミラーリングエージェントが「思い通り-思い通りでない」項目に関して提案エージェントと並び高い評価を得ているが、提案エージェントはミラーリングエージェントに対して有意水準1%で有意に「自由」と評価されている。このことから、提案エージェントはユーザにとって自由かつ思い通りのビーズパターンを創発することが出来るが、ミラーリングエージェントは思い通りではあるがユーザにとって不自由なビーズパターンを創発することが分かる。これはミラーリングエージェントが持つ直前のユーザの記号と同じ記号を出力する挙動の単純さから、ユーザが容易にエージェントの出力を予測できるためと考えられる。

2.5.2 創発されたインタラクション

ここで、本実験環境においてどのようなインタラクションが創発されたか、抜粋して紹介する。なお、記号として用いた「白い丸」、「黄色い丸」、「黒い丸」は便宜上それぞれ「w」、「y」、「b」と表記し、エージェントの出力は太字とする。

単一色



図 2.21: 単一色例

図 2.21 に黒い丸のみで構成されたインタラクション例を示す。文字で表現すると (b,b,b,b,b,b,b,b,b,b,b,b,b,b,b,b) である。このような単一色となるインタラクションは、ミラーリングエージェントとのインタラクションで多く観測出来るものだが、提案エージェントとの学習インタラクションにおいても観測している。そして提案エージェントとの学習インタラクション時には実験参加者から高い報酬を付与されていた。これは本インタラクション環境では最終的に出来上がる図形から数珠を連想しやすいことが影響していると考えられる。

数珠のような並び

図 2.22, 図 2.23, 図 2.24 により数珠を意識したと考えられるインタラクション例を示す。図 2.22, 図 2.23, 図 2.24 はそれぞれ (w,w,b,y,b,w,w,w) (w,w,w,b,w,w,w,w),



図 2.22: 数珠のような並び例 1

(b,w,w,y,w,w,w,w)である．このような数珠を意識したと考えられるインタラクションは，まずメインの色として白か黒が選ばれ，残りの色を一つか二つ入れるといった形式が見られた．特に図 2.23 のような白をメインに黒を間に置いたもの，あるいは黒をメインに白を置いたものがよく見られた．また，図 2.24 のような図 2.23 を回転した形式のインタラクションも見られた．興味深いことに，図 2.23 および図 2.24 のインタラクションが一人の実験参加者から観測できたことである．実験参加者の真の狙いが例えば「数珠のように白を多めに黒を挟んだもの」であったとしても，本提案手法では直接その狙いを考慮してインタラクションを創発することは出来ない．しかし，パターンの解析により「白を多めに，稀に黒を出力」することが実験参加者から高い報酬付与が期待出来ると学習できたことによって，図 2.23，図 2.24 のような回転図形にも対応できたことは提案手法の有効性を示せたと考える．

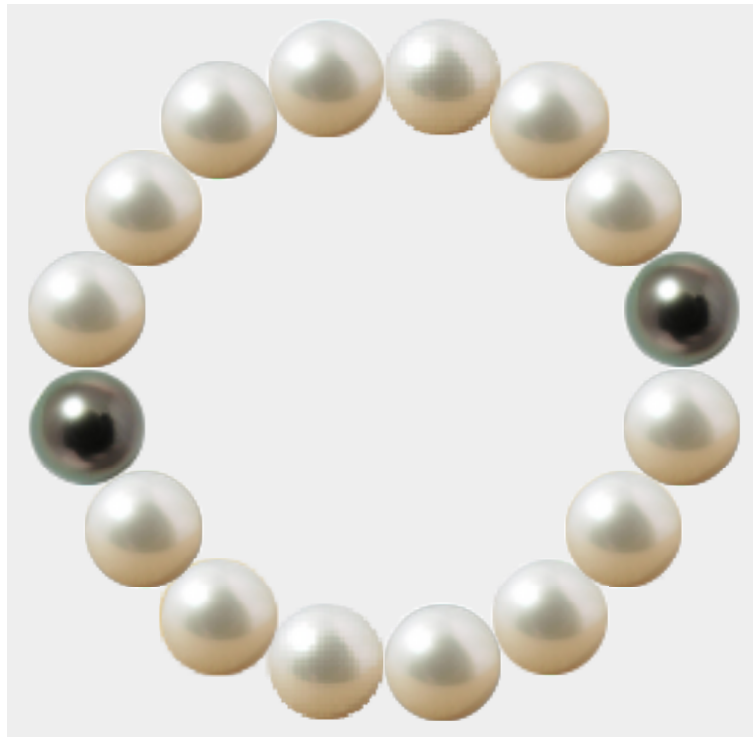


図 2.23: 数珠のような並び例 2

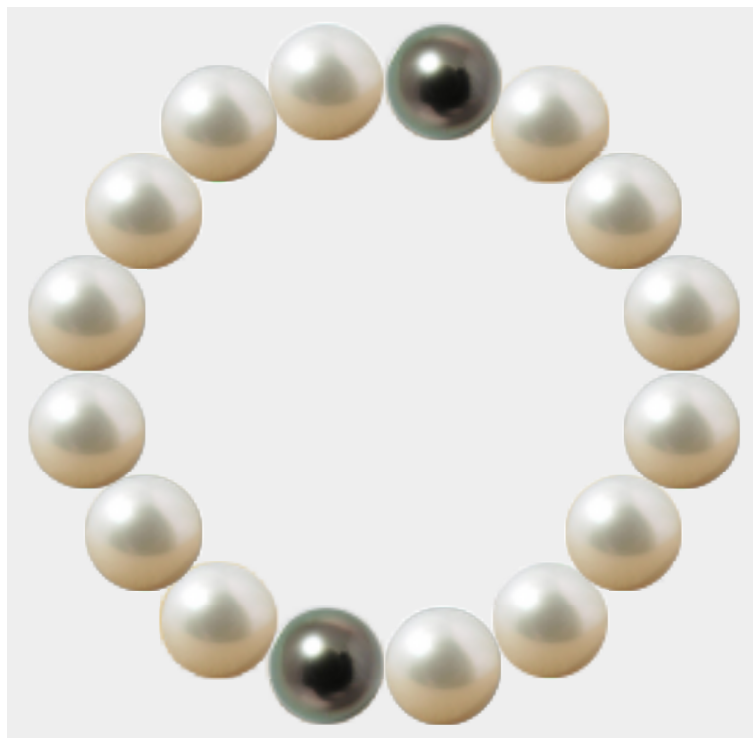


図 2.24: 数珠のような並び例 3

3色をバランスよく配置



図 2.25: 3色をバランスよく配置した例 1

図 2.25, 図 2.26, 図 2.27 に 3 色のバランスを意識したと考えられるインタラクション例を示す。図 2.25, 図 2.26, 図 2.27 はそれぞれ (b, w, y, b, w, y, b, w) (b, y, y, w, w, w, b, b) , (b, w, y, b, b, y, w, b) である。これらのインタラクションは上記の数珠を意識したものと異なり, 3 色の数をバランスよく揃えることを意識していると考えられる。実際, ある実験参加者は「エージェントの出力に合わせて最終的なバランスが整うよう出力を調整した」と実験後のアンケートで回答していた。図 2.25 は黒い丸, 白い丸, 黄色い丸の順に繰り返し出力された結果である。図 2.25 のような 3 色の繰り返しは他にも白い丸, 黄色い丸, 黒い丸の GUI 上の順番に従ったものや, 黒い丸, 黄色い丸, 白い丸と GUI 上の右から左に向かった順番に従ったものも観測した。図 2.26 は色同士がまとまるように意識されて出力されたと考える。即ち最終的に出来上がる図形を強く意識して行われたインタラクションであったと考える。特に, 4 対話目が黒い丸, 黒い丸であることによって 1 対話目に実験参加者が出力した黒い丸と繋がった点から隣接する色を意識している点が見える。図 2.27 は図 2.26 よりもより最終的に出来上がる図形を強く意識して行われたインタラクションであったと考えられる。ある実験参加者はシンメトリーを強く意識しており, 図 2.27 のような上下左右どのように見てもシンメトリーな図形に対して高い報酬を付与していた。提案手法は最終的な図形の形式は学習できないため, このような図形が創発出来たことは 1 対話目の実験参加者の出力である黒い丸と, 3 対話目の出力である黒い丸では次に出力すべき記号が異なることを学習できていたと考える。なお, この図 2.27 のインタラクションに関しては次節にてより詳細に検証する。



図 2.26: 3色をバランスよく配置した例 2



図 2.27: 3色をバランスよく配置した例 3

2.5.3 ユーザ報酬付与傾向の検証

ここで、提案エージェントがユーザの報酬付与傾向を反映出来たかどうか、インタラクションの履歴を用いて考察する．典型例として図 2.20 の評価平均に最も近い感性評価を与えた実験参加者 B を取り上げる．

表 2.10 に実験参加者 B が学習時に最高報酬を与えたインタラクション系列 D_B ，および D_B のユーザの出力を提案エージェントと全体傾向エージェントに与えた場合の創発確率が最も高いインタラクション系列をそれぞれ示す．実験参加者 B が学習時に最高報酬を与えたインタラクション系列 D_B の GUI 上の表示は図 2.27 に示したものと同一である．表 2.11 に実験参加者 B の感性評価結果を示す．また，表 2.12 と表 2.13 に提案エージェントと全体傾向エージェントがそれぞれ獲得したインタラクションルール集合を示す．なお，表中の w, y, b はそれぞれ出力記号の「白い丸」, 「黄色い丸」, 「黒い丸」であり．エージェントの出力に相当する部分は太文字で表記する．また， $F(R)$ は式 2.4 と同様に対話ルール R が持つ報酬履歴に含まれる報酬値の平均値に尤度を積算した値であるが，報酬履歴が存在しない対話ルールの $F(R)$ は ϕ と表記した．

表 2.10: 実験参加者 B のインタラクション系列

学習時に最も高い報酬を得た インタラクション系列 D_B	D_B と同一のユーザ出力を与えた場合に 創発確率が最も高いインタラクション系列	
	提案エージェント	全体傾向エージェント
b,w,y,b,b,y,w,b	b,w,y,b,b,y,w,b	b,b,y,y,b,b,w,b

表 2.11: 実験参加者 B の感性評価結果

評価項目	提案 エージェント	全体傾向 エージェント	ミラーリング エージェント	ランダム エージェント
良い-悪い	1	-1	-1	-1
親しみやすい-親みにくい	1	-1	1	-1
面白かった-つまらなかった	1	1	-2	-1
自由だった-不自由だった	1	1	-2	1
思い通りだった-思い通りでなかった	1	-1	-1	-2

表 2.12: 提案エージェントのインタラクションルール集合 R (表 2.10 の系列に関連する部分)

インタラクションルール R							
1-gram model		2-gram model		3-gram model		4-gram model	
対話系列	$F(R)$	対話系列	$F(R)$	対話系列	$F(R)$	対話系列	$F(R)$
b,w	0.6649	b,w,y,b	1.5958	b,w,y,b,b,y	0.9575	b,w,y,b,b,y,w,b	5
y,b	1.5032	y,b,b,y	1.4361	y,b,b,y,w,b	2.5	b,b,y,y,b,b,w,b	ϕ
b,y	0.1794	b,y,w,b	0.6678	b,b,y,y,b,b	ϕ		
w,b	0.6258	b,b,y,y	ϕ	y,y,b,b,w,b	ϕ		
b,b	-0.2261	y,y,b,b	ϕ				
y,y	ϕ	b,b,w,b	-0.3333				

表 2.10 より，ユーザの出力が D_B と同一の時，提案エージェントが創発する確率が最も高いインタラクション系列は D_B と同一であることが分かる．それに比べ，全体傾向

表 2.13: 全体傾向エージェントのインタラクションルール集合 R (表 2.10 の系列に関連する部分)

インタラクションルール R							
1-gram model		2-gram model		3-gram model		4-gram model	
対話系列	$F(R)$	対話系列	$F(R)$	対話系列	$F(R)$	対話系列	$F(R)$
b,w	0.1710	b,w,y,b	0.2713	b,w,y,b,b,y	0.9575	b,w,y,b,b,y,w,b	5
y,b	0.1026	y,b,b,y	0.3025	y,b,b,y,w,b	3.1915	b,b,y,y,b,b,w,b	3
b,y	0.1838	b,y,w,b	0.7658	b,b,y,y,b,b	0.1995		
w,b	0.2384	b,b,y,y	0.2736	y,y,b,b,w,b	0.0278		
b,b	0.2356	y,y,b,b	0.1407				
y,y	-0.1268	b,b,w,b	0.4091				

エージェントではユーザの出力が同一であるにも関わらず、創発する確率が最も高いインタラクション系列は D_B と異なっていることが分かる。また表 2.12, 表 2.13 の結果から、 D_B (b,w,y,b,b,y,w,b) の創発確率は提案エージェントの場合約 27% であるのに対し、全体傾向エージェントの場合約 19% であり、ランダムエージェントの場合約 1.2% である。このことから、提案エージェントは全体傾向エージェントとランダムエージェントに対して、高い確率で D_B を創発することが可能であることが分かる。以上から、提案エージェントは確率による意外性を持ちつつ実験参加者 B の傾向を反映したインタラクションを創発できていると考えられる。

ここで、表 2.12 と表 2.13 を比較すると、1-gram model と 2-gram model の $F(R)$ が大きく異なることが分かる。また、全体傾向エージェントのインタラクションルールには存在するが、実験参加者 B のインタラクションルールには報酬履歴が存在しない対話ルールがあることも分かる。インタラクション系列の創発確率は式 2.2~2.6 により算出されるため関与する対話ルール R の $F(R)$ によって決まると言ってもよい。実験参加者 B の場合、表 2.4 の (y,b) と (b,w,y,b) と (y,b,b,y) の $F(R)$ が表 2.12 では大きく増加しており、他の対話系列とも明確に差が出ている。これらの差が実験参加者 B の報酬付与傾向として、創発されたインタラクション系列の違いを生んだと考えられる。

表 2.14: ピアソン相関係数

実験参加者	提案エージェント	全体傾向エージェント
A	0.796	0.661
B	0.905	0.358
C	0.721	0.583
D	0.920	-0.153
E	0.865	-0.009
F	0.772	0.575
G	0.866	0.708
H	0.806	0.498
I	0.832	0.151
J	0.828	0.695
K	0.595	0.455
L	0.849	0.835
M	0.823	0.818
all	0.662	0.572

ユーザの傾向はインタラクシヨンルール中の対話ルール R から算出される $F(R)$ の傾向として表現されると考える．学習時に実験参加者から報酬を与えられた系列に対して，その系列の創発に関与する対話ルール R の $F(R)$ を求め，実験参加者が与えた報酬値 U との相関分析を行った．分析には実験参加者 13 名全員の学習データを用いて，ピアソン相関係数を求めた．表 2.14 に実験参加者 13 名の個々のピアソン相関係数および全員のデータを使用したピアソン相関係数を示す．

分析の結果，提案エージェントにおける全員のデータを使用した相関係数は 0.662 であり，ユーザ報酬値 U と $F(R)$ の高い相関が確認された．一方で，同様の分析を全体傾向エージェントに対して行ったところ相関係数は 0.572 であり，提案エージェントと比べて相関係数は低かった．各実験参加者のデータにおいてもそれぞれ相関係数を求めたところ，全実験参加者において全体傾向エージェントにくらべ提案エージェントの相関係数が高いことがわかった．さらに提案エージェントにおける相関係数は 13 名中 12 名が 0.6 を超えているが，全体傾向エージェントにおける相関係数は 13 名中 5 名のみが 0.6 を超えていることがわかった．以上から，提案エージェントはユーザ報酬値と相関の高いインタラクシヨンルールを生成することができたことが分かる．

2.5.4 記号毎の出力傾向

表 2.15: 記号毎の出力回数（実験参加者）

実験参加者の出力	合計（全実験参加者）	割合
白い丸	366	35.2%
黄色い丸	197	18.9%
黒い丸	477	45.9%
全記号	1040	

表 2.15 に記号毎の全実験参加者合計出力回数を示す．表 2.15 より，実験参加者自身の出力は黒い丸が最も多く，黄色い丸が最も少なかったことがわかる．2.4.5 節における「短音」を用いた実験時の記号と比べ，記号間の出力数に大きな偏りが見られた．これに関しては，実験参加者の好みの偏り以上に要因があったと考える．要因の一つとして最終的に出来上がる絵が「数珠」のように見えることが挙げられる．2.5.2 節で紹介したように数珠を意識したインタラクションも多数観測している．ある実験参加者は報酬付与の方針として「数珠っぽくなるようにしたかった」と述べ，特に実験参加者自身は黒い丸を出力していた．数珠は黒色のみならず，多種多様なものが存在し，かつ色に対する厳密な決まりも無い．しかし，一般的に葬式は黒を基調としているため，数珠の色も黒である印象が強く，表 2.15 のような出力回数の偏りが生じたと考える．

表 2.16: 記号毎の出力回数（提案エージェント）

提案エージェントの出力	合計（全実験参加者）	割合
白い丸	311	29.9%
黄色い丸	309	29.7%
黒い丸	420	40.4%
全記号	1040	

表 2.16 に提案エージェントにおける記号毎の合計出力回数を示す．表 2.16 より，エージェントの出力は実験参加者と同じく黒い丸が最も多く，黄色い丸が最も少なかったことがわかる．しかし，表 2.15 に示した実験参加者の出力傾向ほど極端な偏りにはなっていない．エージェントの出力傾向は実験参加者からの報酬付与によって決まることから，本実験の参加者は自身の出力には黒い丸を選びつつも，エージェントの出力が黒い丸となること，あるいは黄色い丸を選ばないことには強いこだわりがなかったことが伺える．

2.6 おわりに

本章では、ユーザの好むインタラクションを創発するためのユーザ報酬付与傾向を動的に学習する手法として All-Combinatorial N-gram (ACN) を提案した。ACN ではインタラクション系列の分割および、インタラクションルールを生成、生成されたインタラクションルールに基づきインタラクションを創発する。GUI を用いた感性評価実験によって、提案手法を搭載したエージェントは他のエージェントと比べ、ユーザにポジティブな印象を与えるインタラクションを創発することを確認した。感性評価実験は「音」および「絵」の記号を用いて行い、記号の差異が提案手法の有効性に影響を与えないことを確認した。創発されたインタラクションについても特徴的なものを取り上げ、それらの創発に提案手法が有効に働いていることを確認した。また、インタラクションの長さの違いによる検証実験により一回のインタラクションが長い時の評価値の上がり方を検証し、最適な長さを考察した。また、提案エージェントのインタラクションの創発に関して、確率的に行なっていることの有効性を決定的に出力するエージェントとの比較によって検証した。さらに被験者の報酬のずれを考慮し、インタラクションルール内の対話ルールに対して確率密度による重み付けをおこない方式に対して、確率密度を考慮しない方式との比較実験を行い有効性を検証した。

インタラクションルールの検証により、ユーザ報酬と相関の高いルールであること、ユーザ報酬が高いインタラクションの創発確率が比較対象よりも高いことを各記号毎に確認した。しかし、被験者の内部に存在すると考えられる真のインタラクションルールと提案エージェントが生成したインタラクションルールとの差異は検証出来ておらず、最終的に出力された系列ではなく、ルールそのものに対する主観評価実験を行う必要があると考える。

また、本手法は出現パターンの学習であるが、現段階ではユーザとエージェントが交互に記号を出力するだけの原始的なインタラクションしか扱うことが出来ない。実環境における様々なインタラクションに応用するためにはモデルを改良しより高度なインタラクションを扱える必要がある。さらに、2.4.6 節で確認したようにインタラクション系列が長い場合は学習回数が 20 回では足りない可能性が高く、学習回数に制限がある実験形式では長いインタラクション系列に対する提案手法の有効性が確認できていない。インタラクション系列長が固定である点はインタラクションの創発性も減少させている。2.4.3 節および 2.5.2 節で確認したような特徴的なインタラクションは明確なパターンであったために観測が容易だったが、インタラクションの創発性とはインタラクションを行う両者が予想すら出来ないものが出来上がることであり、本実験にて創発されたインタラクションでは創発性の確認が十分に行なえたとはいえないと考える。

このように、提案手法の有効性を検証するためには検証に用いるインタラクションの改良および実験設定の見直しが必要である。次章ではこれらを考慮したインタラクションモデル、および実験について述べる。

第3章

ACN およびインタラクションモデルの 拡張

3.1 はじめに

2章におけるインタラクション環境は非常に単純であり、All-Combinatorial N-gram (ACN) の挙動を検証する際には適していたが、その応用性は十分に議論できなかった。本章ではACNの応用性への議論の第一歩として「ボールを使ったやりとり遊び」を題材とする。「ボールを使ったやりとり遊び」とは乳児と養育者間でよく行われる「やりとり遊び」の一種であり、本稿では特に「ユーザとエージェントがボールを両者間でやりとりする遊び」とする。具体的にはボールを相手に向かって差し出す、差し出されたボールを受け取る、ボールを相手に向かって投げる、などの行動をユーザとエージェント間で行うインタラクションである。そのため本章におけるエージェントとは、ボールを掴むことが可能であり、エージェントがボールを掴んでいることをユーザが容易に認識出来るよう、擬人化エージェントを採用する。

図3.1に本章におけるエージェントの外観を示す。本章ではPlayer Charactor とエージェントが手先を動かし、ボールを介してやりとりすることを「ボールを使ったやりとり遊び」と呼称する。

本章では「ボールを使ったやりとり遊び」を扱えるよう、より一般化したACNを提案し、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実験参加者による感性評価を基に有効性を示す。

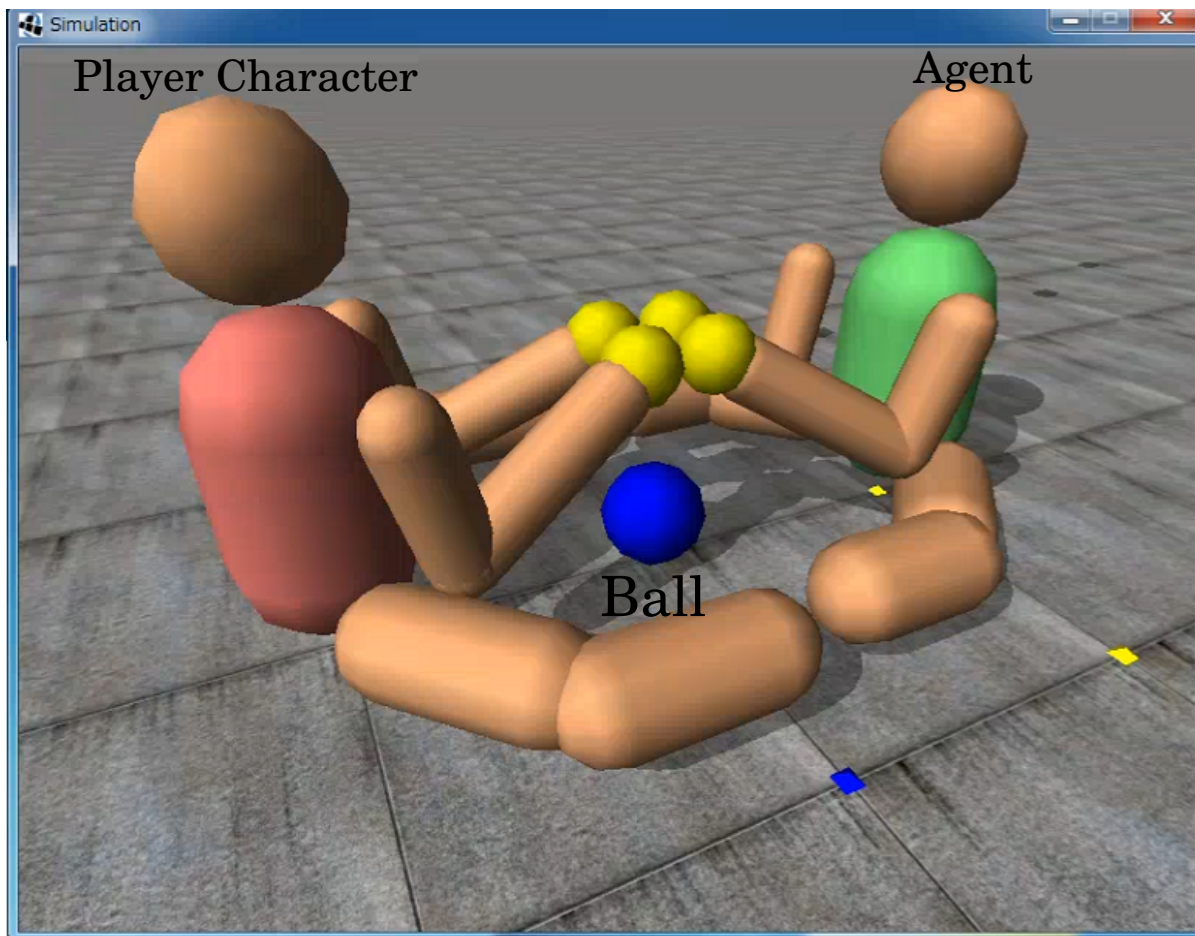


図 3.1: エージェントの外観

3.2 本章におけるインタラクション

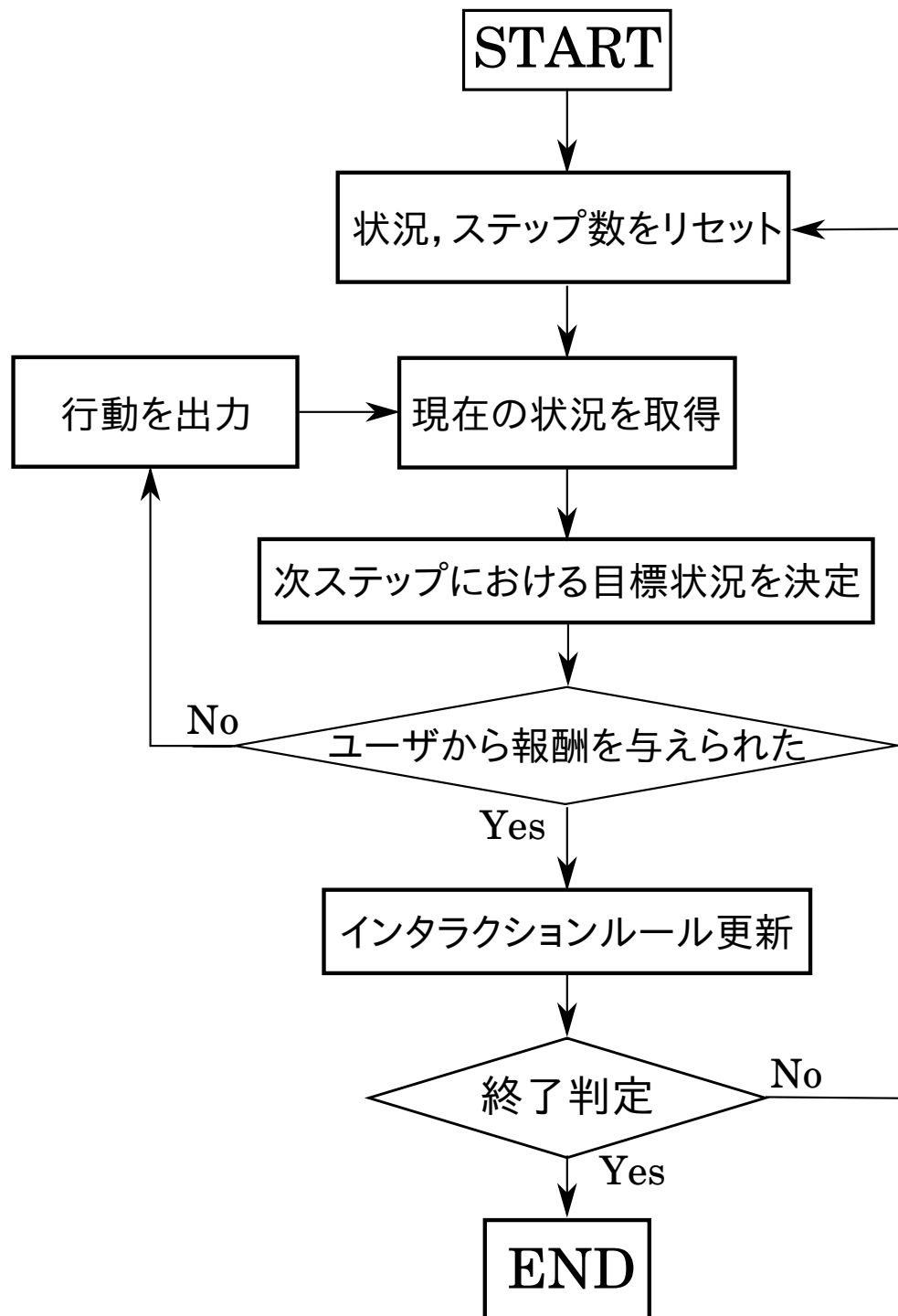


図 3.2: エージェントの行動決定プロセス

本章ではユーザがいつでも報酬付与を行えるインタラクションを題材とする。図 3.2 に本章におけるエージェントの行動決定プロセスを示す。なお、本章ではインタラクションにかかわる情報を「状況」と呼び、エージェントは1ステップごとに状況を取得可能

であるとする。エージェントは現在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し、目標状況に移行するための行動を出力する。目標状況の決定方法に関しては3.3.2にて述べる。

例として、「ボールを使ったやりとり遊び」の場合、ユーザとエージェントの姿勢（手先位置）やボールの状態などが「状況」であり、エージェントは目標状況決定後、目標状況内のエージェントの姿勢に一致するように自身の姿勢を変化させる。インタラクションルールの更新はユーザが与えた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。初期状況からユーザが報酬を付与するまでの状況の系列を「インタラクション系列」と呼ぶ。ユーザとエージェントは非同期的に行動を決定し、状況はリアルタイムに変化するものとする。

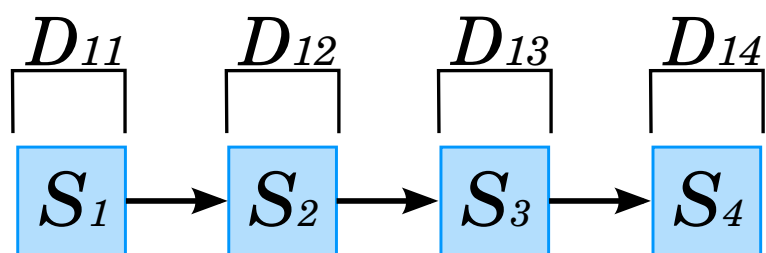
2章ではインタラクションに以下の制約を設けた。

- ユーザ、エージェントのどちらもあらかじめ設定された出力記号集合の中から記号を選び出力する
- ユーザとエージェントは交互に記号を出力する
- ユーザとエージェントは同じ回数だけ記号を出力する

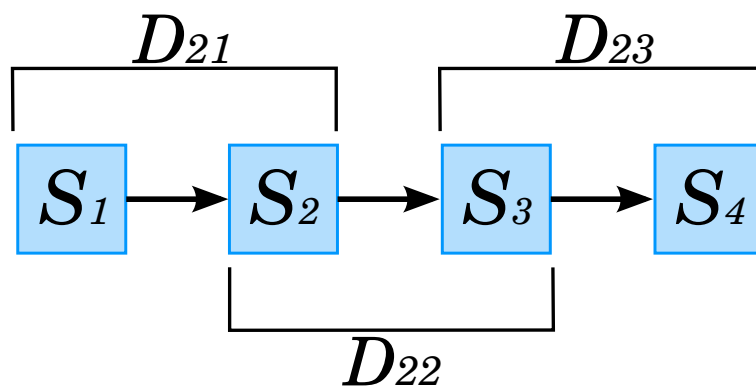
上記制約と比較して、本章で扱うインタラクションは以下のようなになる。

- ユーザ、エージェントはあらかじめ設定された行動集合の中から行動を選び出力する
- ユーザとエージェントは非同期的に行動を決定する
- ユーザとエージェントの出力回数に制限は無い

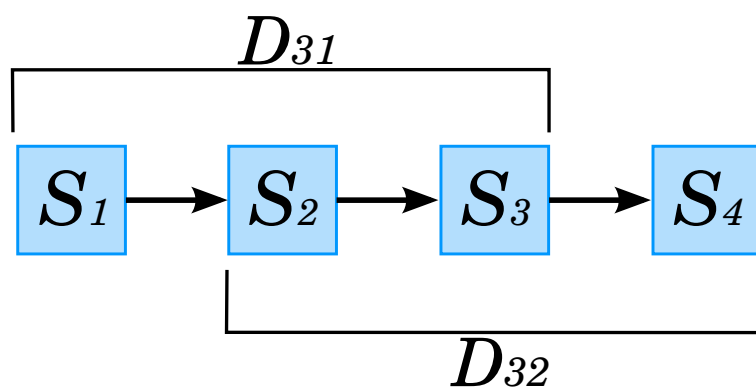
3.3 ACN の改良



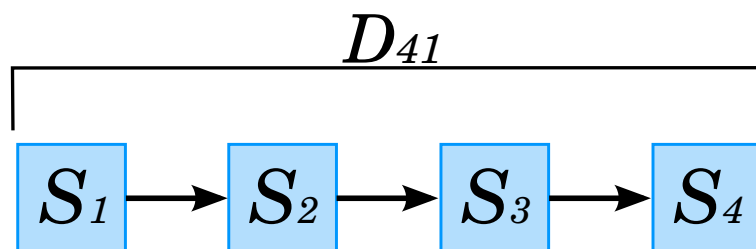
(a) 1-gram



(b) 2-gram



(c) 3-gram



(d) 4-gram

図 3.3: ACN による状況系列分割例 (4 ステップ目までの場合)

2.3 節にて提案して ACN は 2 章におけるインタラクションに特化していたため，本章ではまず ACN を改良し，抽象度を上げることとする．まず，本章では 2 章とは異なり

「対話」ではなく「状況」の系列を ACN によって学習する。 S_t を t ステップ目の状況であり、4 ステップ目までの状況系列は図 3.3 のように分割される。このように分割された後、ユーザから付与された報酬を基にそれぞれ学習する。

なお、本章ではインタラクション系列の部分系列である、報酬付与時点から遡って T_{max} 個の状況までの系列に対して報酬を付与するものとする。これを「報酬付与対象系列」と呼ぶ。インタラクション系列長が T_{max} 未満の場合、報酬付与対象系列はインタラクション系列と等しい。本章では報酬付与対象系列を ACN によって分割し、分割後の系列を「パターン系列 D_{ij} 」と呼ぶ。 i ($1 \leq i \leq T_{max}$) はパターン系列の長さを表し、 j は i -gram モデルごとのパターン系列を表す添字である。すなわちパターン系列は報酬付与対象系列の部分系列であり、系列長が $1 \sim T_{max}$ である「状況」の系列である。 T_{max} はインタラクションの種類、状況に含む情報量などに応じて決定するのが望ましい。本章における具体的な値は 3.4 節にて述べる。

3.3.1 インタラクションルールの更新

本章ではACN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 R と呼ぶ． R は今までの経験を N-gram ルールとして保持しており，ユーザが報酬を与える度に更新される．本章における N-gram ルールはパターンルールの長さごとの集合であり， 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．パターンルール R_{ij} は，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．

パターン系列 D_{ij} のある報酬値 P 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 3.1, 3.2 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．

$$P_L = \frac{U \times A_{ij}}{(1 + e^{T - \|D_{ij}\| - T_{max}/2})} \cdot \quad (3.1)$$

$$P_D = \frac{U \times A_{ij}}{(1 + e^{T - K_{ij} - T_{max}/2})} \cdot \quad (3.2)$$

ここで i は N の値を表し， j は i -gram モデル内のパターン系列の識別子， U はユーザが与えた報酬値， A_{ij} は報酬付与対象系列内に D_{ij} が出現した個数， T は報酬付与対象系列長， $\|D_{ij}\|$ は D_{ij} のパターン系列長， T_{max} は最大パターン系列長， K_{ij} は報酬付与時から D_{ij} の末尾までの距離（ステップ数）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，本章における具体的な値および報酬付与方法は 3.4 にて述べる．式 3.1 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に対して大きな値となり，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 3.2 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に対して大きな値となり，タイミングに重きをおいた学習を行うことができる．

本章では式 3.1, 3.2 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 3.3 に本章で使用する報酬値 P_{DL} の算出式を示す．

$$P_{DL} = \frac{U \times A_{ij}}{(1 + e^{T - \|D_{ij}\| - T_{max}/2})(1 + e^{T - K_{ij} - T_{max}/2})} \cdot \quad (3.3)$$

3.3.2 目標状況の決定

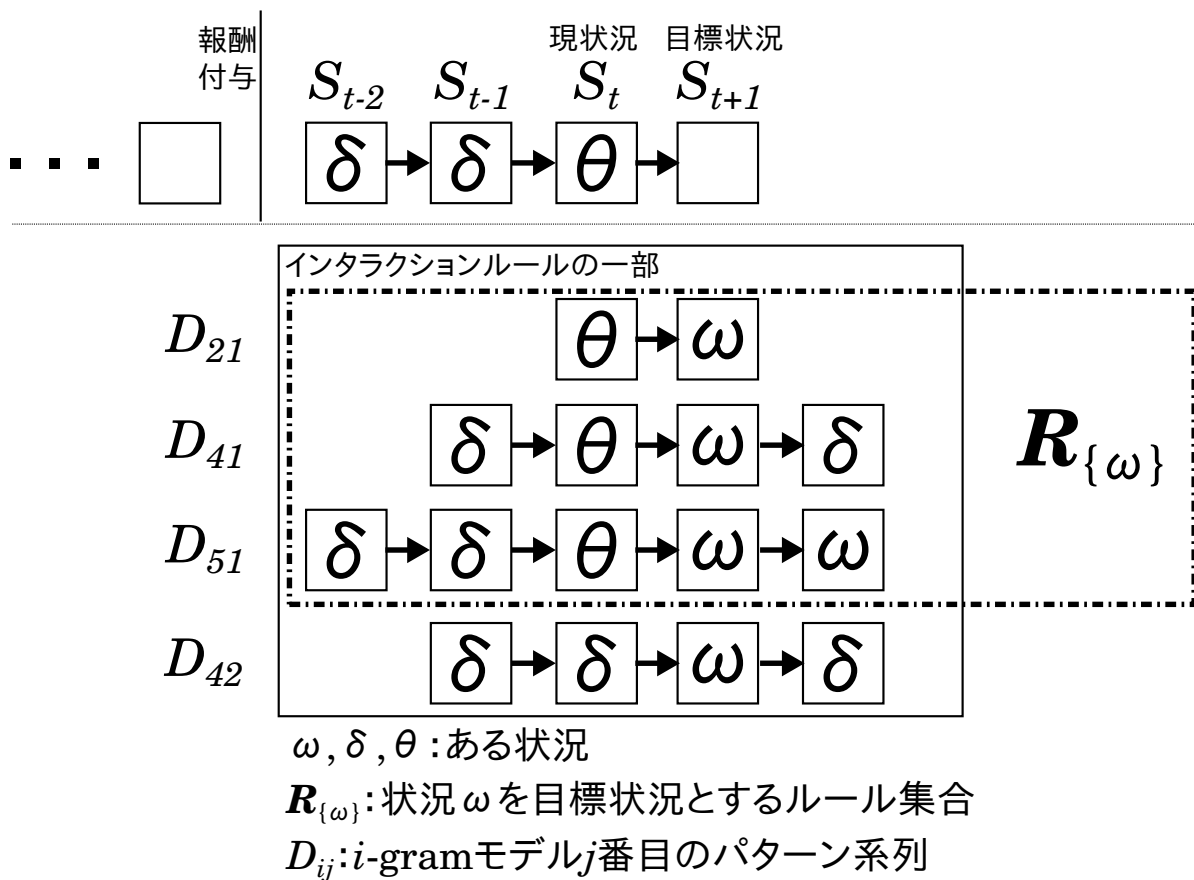


図 3.4: α_k 以前に状況が存在する D_{ij} の場合の $R_{\{\alpha_k\}}$ 選出例

エージェントはユーザからの報酬を最大化するような状況を目指とする。次ステップにおけるエージェントの目標状況は獲得したインタラクションルール R 及び、現在のステップから遡って T_{max} 個までの状況系列によって決定される。なお、ステップ数が T_{max} 未済の場合は初期状況から現状況までの状況系列が目標状況の決定に用いられる。目標状況は以下の手順で決定する。

エージェントが学習済みの状況の集合を α とし、目標状況の候補を $\alpha_k \in \alpha$ とすると、エージェントは α_k に対する総報酬予測値 E_{α_k} を以下の手順で算出する。

1. パターン系列 D_{ij} に状況 α_k を含むパターンルール R_{ij} の内、以下の条件を満たすものをパターンルールの集合 $R_{\{\alpha_k\}}$ として求める。なお、 D_{ij} はパターンルール R_{ij} の持つパターン系列を、* は任意の状況を表す。

- α_k 以前に状況が存在しない D_{ij} の場合

$$R_{\{\alpha_k\}} = \{\forall R_{ij} \in R \mid (\alpha_k, *, \dots, *) = D_{ij}\}. \quad (3.4)$$

- α_k 以前に状況が存在する D_{ij} の場合
 α_k 以前の状況の系列を β , β の系列長を $\|\beta\|$ とすると , 初期状況 S_0 から現状 S_t までの状況の系列は以下のように表現できる .

$$S_0, \dots, S_{t-\|\beta\|+1}, \dots, S_t \quad (3.5)$$

また $R_{\{\alpha_k\}}$ は以下の条件のもと求める .

$$R_{\{\alpha_k\}} = \{ \forall R_{ij} \in \mathbf{R} \mid (S_{t-\|\beta\|+1}, \dots, S_t) = \beta \wedge (\beta, \alpha_k, *, \dots, *) = D_{ij} \} . \quad (3.6)$$

図 3.4 に前回の報酬付与後 , 状況が δ, δ, θ と遷移した現状における $\alpha_k = \omega$ の場合の $R_{\{\alpha_k\}}$ 選出例を示す . 図 3.4 では , 上記条件に基づきパターン系列 D_{21} , D_{41} , D_{51} を持つパターンルール R_{21} , R_{41} , R_{51} を $R_{\{\omega\}}$ の要素として選出している . D_{42} は $\beta = (\delta, \delta)$ が現状までの系列 ($S_{t-1}, S_t = (\delta, \theta)$) と一致しないため選出されない .

2. α_k の総報酬予測値 E_{α_k} を下式にて求める .

$$E_{\alpha_k} = \sum_{R \in R_{\{\alpha_k\}}} F(R) . \quad (3.7)$$

$$F(R) = \begin{cases} \mu(R) \times \frac{1}{\sqrt{2\pi}\sigma(R)} & (\sigma(R) \neq 0) , \\ \mu(R) & (\sigma(R) = 0) . \end{cases} \quad (3.8)$$

ここで $\mu(R)$ は R が持つ報酬履歴に含まれる報酬値の平均値を , $\sigma(R)$ は標準偏差を表す . なお式 3.8 は正規分布の確率密度関数式 3.9 を参考に , 平均値の確率密度 $f(\mu(R))$ を尤度として $\mu(R)$ に積算した .

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) . \quad (3.9)$$

以上の手順を状況集合 α 内の全要素に対して計算し , 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 OP_k を算出する . そして OP_k に基づいて確率的に目標状況を決定する .

$$OP_k = \frac{E_{\alpha_k}}{\sum_{\alpha_k \in \alpha} E_{\alpha_k}} \times 100 . \quad (3.10)$$

エージェン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 . なお , 抽出されたパターンルールが一つもない場合 ($R_{\{\alpha_k\}} = \phi$ for $\forall \alpha_k \in \alpha$) は目標状況を決定できないため , ランダムに行動を出力する .

3.4 感性評価実験

人 - エージェント間のインタラクションにおいて本章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。

3.4.1 計算機実装

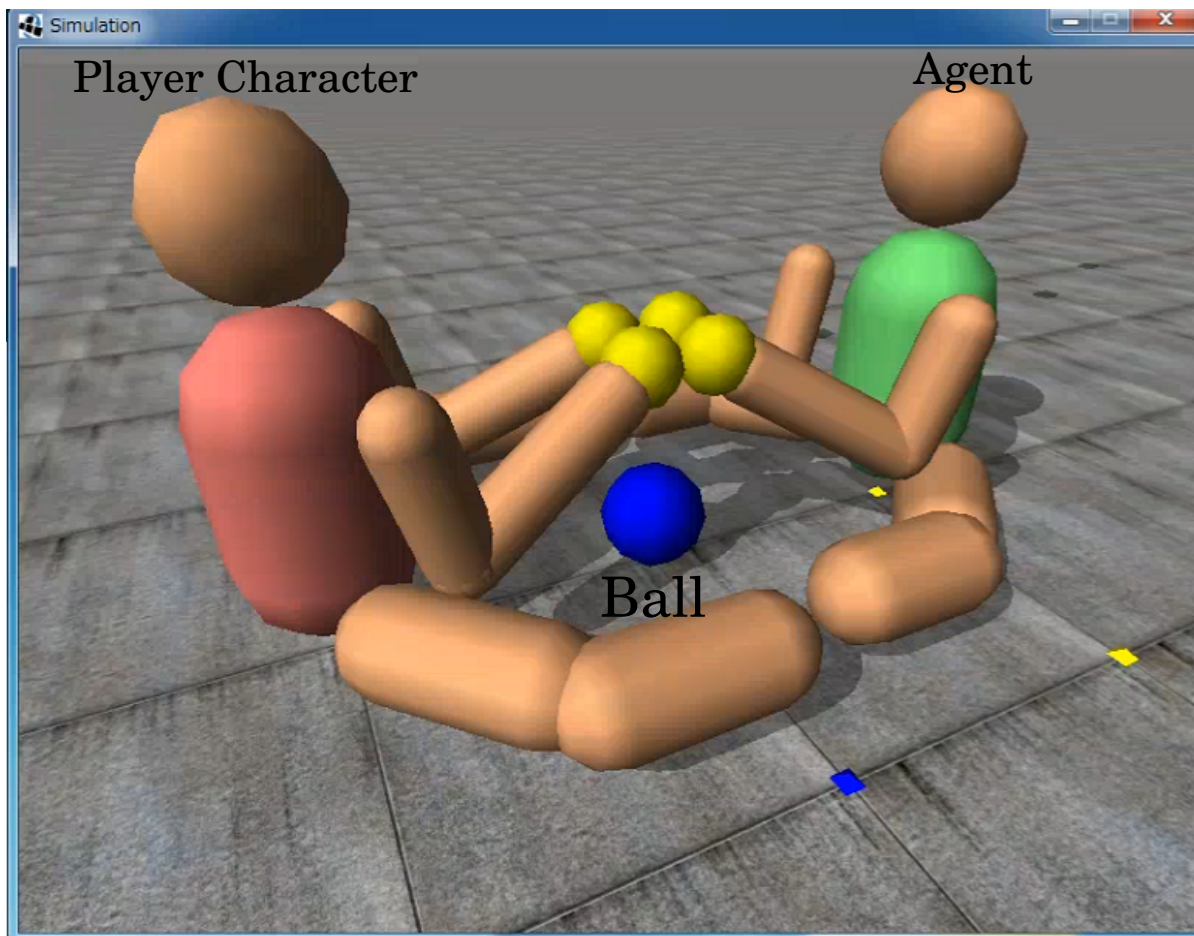


図 3.5: シミュレータ上のキャラクターとボールの外観

本章では人 - エージェント間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 [48, 40] が使用したものを参考に用意した。栗山らは「親と子のボールを使ったやりとり遊び」をモチーフとした環境をインターフェースとして構築し、インタラクション実験の中で多彩なやりとりが人 - ロボット間で創発されたこと、および栗山らが想定していなかったやりとりも創発されたことを確認している [40]。我々は栗山らのインターフェースを参考にインターフェースを構築することで、2章と比較してより複雑で多様なインタラクションが創発できる環境において提案手法の有効性を検証できると考えた。なお、構築したインターフェースは [48, 40] を参考に我々が作成したものであり、栗山らが用いたものと厳密には一致しない。以下に実装の詳細を述べる。

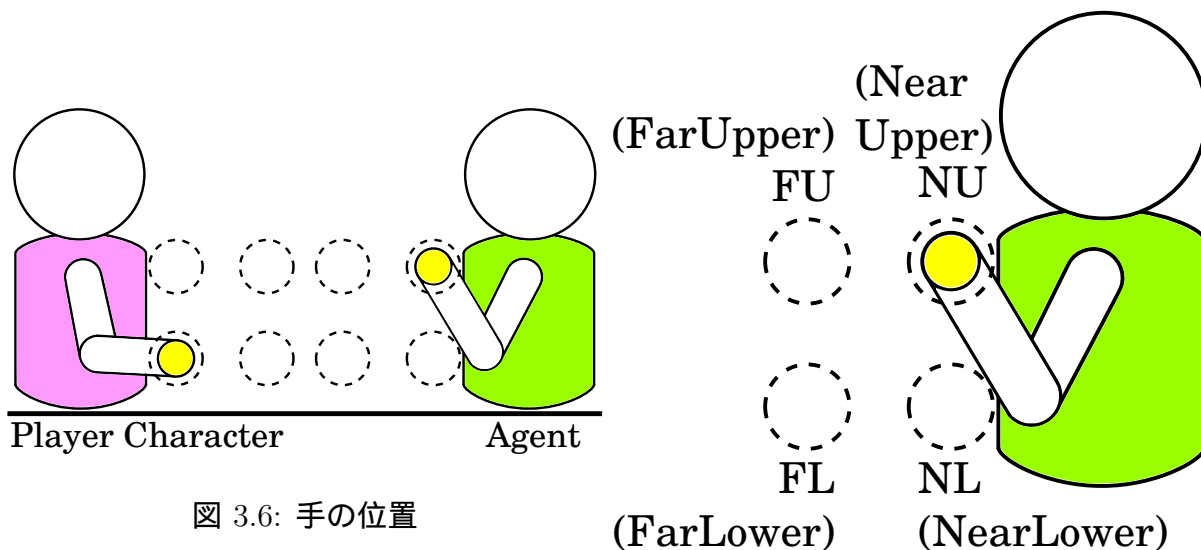


図 3.6: 手の位置

図 3.7: 手を移動できる場所

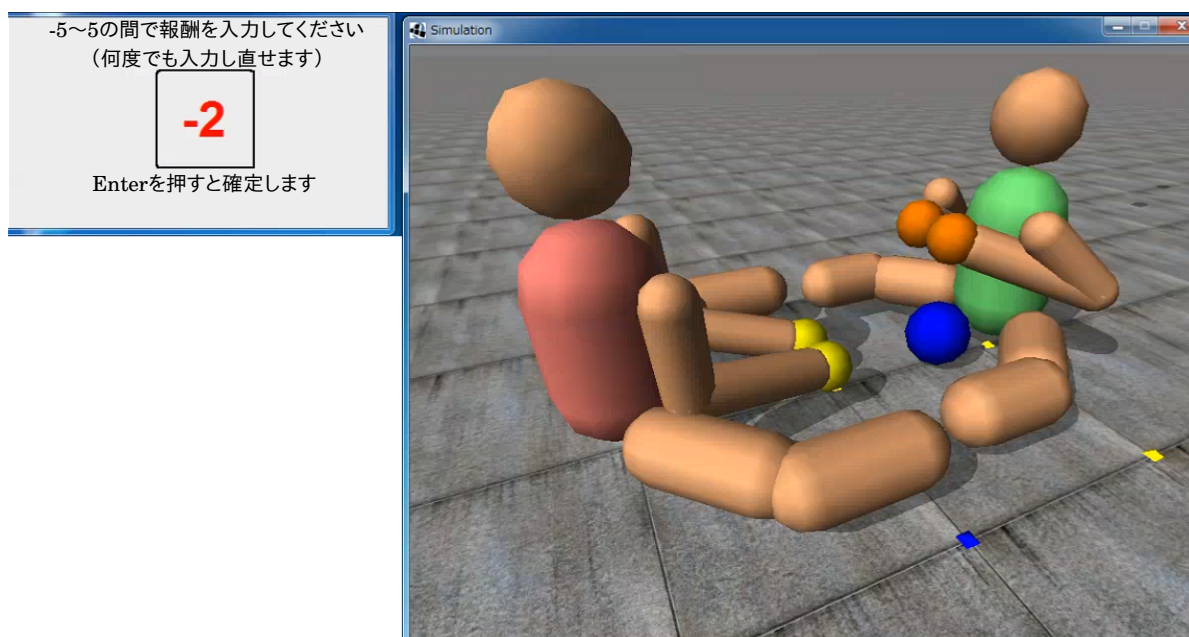


図 3.8: 報酬付与中の様子

ボールを使った遊びに場面を設定し、シミュレータ上に人型キャラクターを二体構築し向き合う形で配置した。本章では実験参加者が操作するキャラクターを「PC (Player Character)」, もう片方を「エージェント」と呼称する。PC とエージェントの間にはボールを配置した。図 3.5 に実験に用いた人型キャラクターおよびボールの外観を示す。なお図 3.5 における人型キャラクターの手先位置およびボールの位置は初期状態であり、インタラクションは毎回この状態から始まる。

PC は実験参加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

キャラクター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．エージェントとPC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．

本章ではPCとエージェント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3.6）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3.7）．初期手先位置は両者ともFUである．なお，本章における「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ないとする．手の握りは「握っている」，「握っていない」の2状態とする．すなわち，PCおよびエージェントが選択できる状況の種類は{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種類である．ボールの状態数は「PCがボールを掴んでいる」，「エージェント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できる．

エージェントは1秒ごとに状況を取得する．取得する状況に含まれる情報は，PCの手の8状態，エージェントの手の8状態，ボールの4状態の3次元であり，計256通りである．エージェントは状況を取得後，1秒後の目標状況を決定し，目標状況を達成するようエージェントの手を操作する．エージェントはインタラクション中，取得した状況を最新8 ($T_{max} = 8$) 個記憶する．よって報酬付与対象系列の最大長は8となり，実験参加者から与えられた報酬は最新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．

身心共に健康な20代男性10名の実験参加者が実験に参加した．事前にPCの操作方を説明し，インタラクションに必要な動きを練習させた上で実験に臨んだ．本実験環境はリアルタイムに動作するものであるため，特に任意のタイミングで任意の場所にPCの手先が移動できること，報酬付与のためのボタン操作をスムーズに行えることを確認できるまで練習させた．本実験ではエージェントとのインタラクションについて「良い」と感じた場合に高い報酬を与えるよう実験参加者に指示する．どのようなインタラクションを「良い」と思うかは実験参加者の主観に任せるものとする．実験参加者からの報酬値 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．図3.8に実験参加者がエージェントに報酬を与えている様子を示す．報酬を与える際はまずシミュレータを一時停止させた上で，数字をキーボードから入力する．シミュレータはボタン一つで停止できるよう設計したため，リアルタイムにインタラクションが進行する中でも任意のタイミングで報酬を与えやすい環境となっている．なお，報酬付与とエージェントの目標状況決定タイミングが衝突した場合は，報酬付与後に更新後のインタラクションルールを用いて新たに目標状況を決定する．

3.4.2 感性評価

実験参加者には5つのエージェントと5分間やりとりさせ、各エージェントとのやりとり終了後に感性評価をしてもらった。順序効果を考慮し、やりとりするエージェントの順番は実験参加者ごとにランダムに決定した。なお、各エージェントとのやりとりごとに5分程度の休憩を挟んだ。評価実験に用いたエージェントを以下に示す。

- 提案エージェント (エージェント DL):
提案手法 (式 3.3) によりユーザ報酬付与傾向を学習するエージェント
- 長さ優先エージェント (エージェント L):
提案手法のうちインタラクションルール更新時に式 3.1 を使用するエージェント
- 近さ優先エージェント (エージェント D):
提案手法のうちインタラクションルール更新時に式 3.2 を使用するエージェント
- ランダムエージェント (エージェント R):
ユーザからの報酬を無視し、ランダムに動くエージェント
- ミラーリングエージェント (エージェント M):
ユーザの動きを真似するエージェント

なお、長さ優先エージェントおよび近さ優先エージェントは本章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 3.3) の有効性の検証のため採用した。それぞれ更新時に式 3.1 (計算時に系列の長さのみを考慮する式) を用いるエージェント、式 3.2 (報酬付与時からの近さのみを考慮する式) を用いるエージェントとなっている。また、ミラーリングエージェントはエージェントの状況取得時における PC 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 PC の手の状態が「NU・握っていない」であれば、ミラーリングエージェントは「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 [45] を用い、以下の形容詞対について7段階評価で行う。実験に用いる形容詞対を以下に示す。

1. 賢い - 愚かな
2. 人間的な - 機械的な
3. 敏感な - 鈍感な
4. 派手な - 地味な
5. 複雑な - 単純な
6. 感じの良い - 感じの悪い
7. 面白い - つまらない
8. 積極的な - 消極的な
9. 好きな - 嫌いな

10. 自然な - 違和感のある

上記形容詞対は、神田ら [49] などのエージェントに抱く印象研究およびSD法についてまとめている書籍 [50]などを参考に感性工学に携わる研究者と協議の末に決定した。また、全エージェントとのインタラクション後に「最も印象が良かったエージェント」および「最も印象が悪かったエージェント」を聴取した。

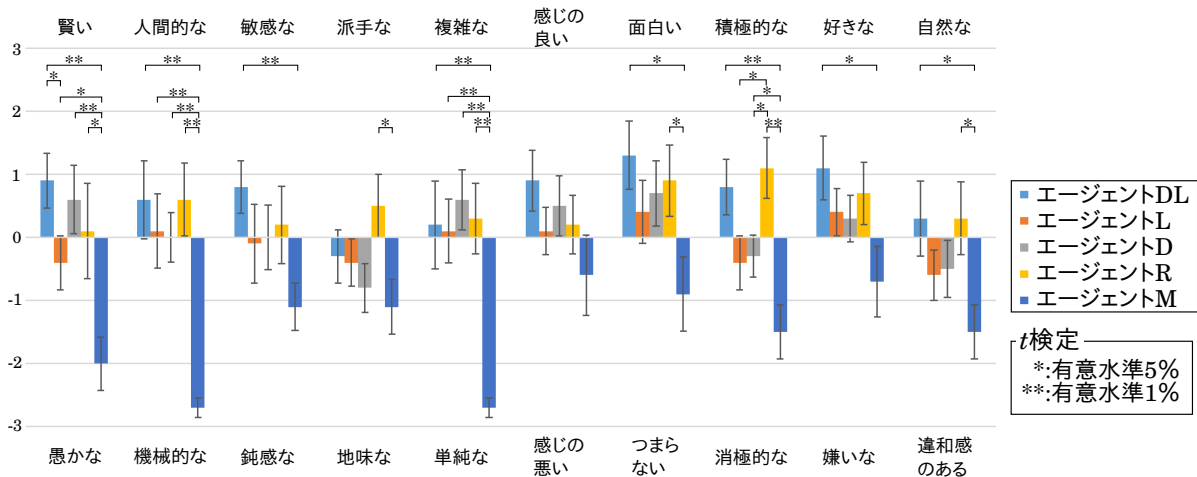


図 3.9: 感性評価実験の結果

図 3.9 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、各エージェントの評価に対して t 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% にてエージェン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 3.9 より、提案エージェント（エージェント DL）は「賢い」「人間的な」「敏感な」「複雑な」「面白い」「積極的な」「好きな」「自然な」でミラーリングエージェント（エージェント M）に対して有意差を確認できたことがわかる（「賢い」： $t = 4.796, p < 0.001$ 。「人間的な」： $t = 5.182, p < 0.001$ 。「敏感な」： $t = 3.376, p = 0.003$ 。「複雑な」： $t = 4.070, p = 0.001$ 。「面白い」： $t = 2.764, p = 0.013$ 。「積極的な」： $t = 3.737, p = 0.002$ 。「好きな」： $t = 2.391, p = 0.028$ 。「自然な」： $t = 2.449, p = 0.025$ ）。ランダムエージェント（エージェント R）に対しては「賢い」「敏感な」「感じの良い」「面白い」「好きな」という形容詞対において提案エージェントの方が高い評価であることを確認した。しかし、 t 検定における有意差は確認できなかった（「賢い」： $t = 0.922, p = 0.372$ 。「敏感な」： $t = 0.812, p = 0.428$ 。「感じの良い」： $t = 1.043, p = 0.311$ 。「面白い」： $t = 0.512, p = 0.615$ 。「好きな」： $t = 0.566, p = 0.579$ ）。

ミラーリングエージェン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、本章におけるインタラクションには向いていないエージェントであったと考える。2章にて行ったインタラクション実験においてはミラーリングエージェントが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エージェントであると考えられる。

表 3.1: 最も印象が良かった・悪かったエージェント

実験参加者	最も良い	最も悪い
A	エージェント DL	エージェント M
B	エージェント D	エージェント M
C	エージェント D	エージェント M
D	エージェント DL	エージェント M
E	エージェント M	エージェント R
F	エージェント D	エージェント R
G	エージェント R	エージェント M
H	エージェント DL	エージェント M
I	エージェント DL	エージェント D
J	エージェント DL	エージェント D

表 3.1 に実験参加者が「最も印象が良かった、または悪かったエージェント」というアンケートに回答した結果を示す。表 3.1 より、提案エージェント（エージェント 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エージェント（エージェント R）との有意な差は確認できなかったものの、実験参加者の中では「最も印象が良かった」エージェントである傾向が見られることから提案エージェントの有効性が示せたと考える。なお、近さ優先エージェント（エージェント D）は提案エージェントについて多く「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、実験参加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、ミラーリングエージェント（エージェント M）が最も多く「最も印象が悪かった」と評価されており、感性評価実験の結果と同等の結果であることを確認した。

3.4.3 エージェントごとの報酬付与頻度の違い

表 3.2: エージェントごとの報酬付与頻度

	報酬付与頻度（全実験参加者平均）
提案エージェント	1.72
長さ優先エージェント	1.3
近さ優先エージェント	1.42
ランダムエージェント	1.58
ミラーリングエージェント	0.88

本章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エージェントに報酬を与えることができる。我々は実験参加者がインタラクション相手のエージェントに対して学習することを期待しているかどうかによって報酬付与頻度¹が変化するのはないかと考え、感性評価実験に用いたエージェントごとの報酬付与頻度の違いに着目した。表 3.2 にエージェントごとの報酬付与頻度における全実験参加者平均を示す。表 3.2 より、ミラーリングエージェントに対する報酬付与頻度が最も低いことがわかる。これはミラーリングエージェン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。ミラーリングエージェントの「ユーザの動きを真似する」仕組みには全実験参加者がインタラクション中に気づいており、ミラーリングエージェントが学習することに期待を持たなくなったと考えられる。なお、提案エージェントに対する報酬付与頻度は最も高かったため、提案エージェントに対しては学習を期待する実験参加者が多かったと考える。

¹各エージェン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

3.4.4 創発されたインタラクション

本章におけるインタラクション環境は2.4節および2.5節で構築した環境よりも複雑かつ多様なインタラクションが創発されることを目的として構築している．本節では感性評価実験のなかで実験参加者 - 提案エージェント間で実際に創発されたインタラクションを例に本環境について考察する．以下に本実験にて創発されたインタラクションの例を示す．

「ボールの受け渡し」

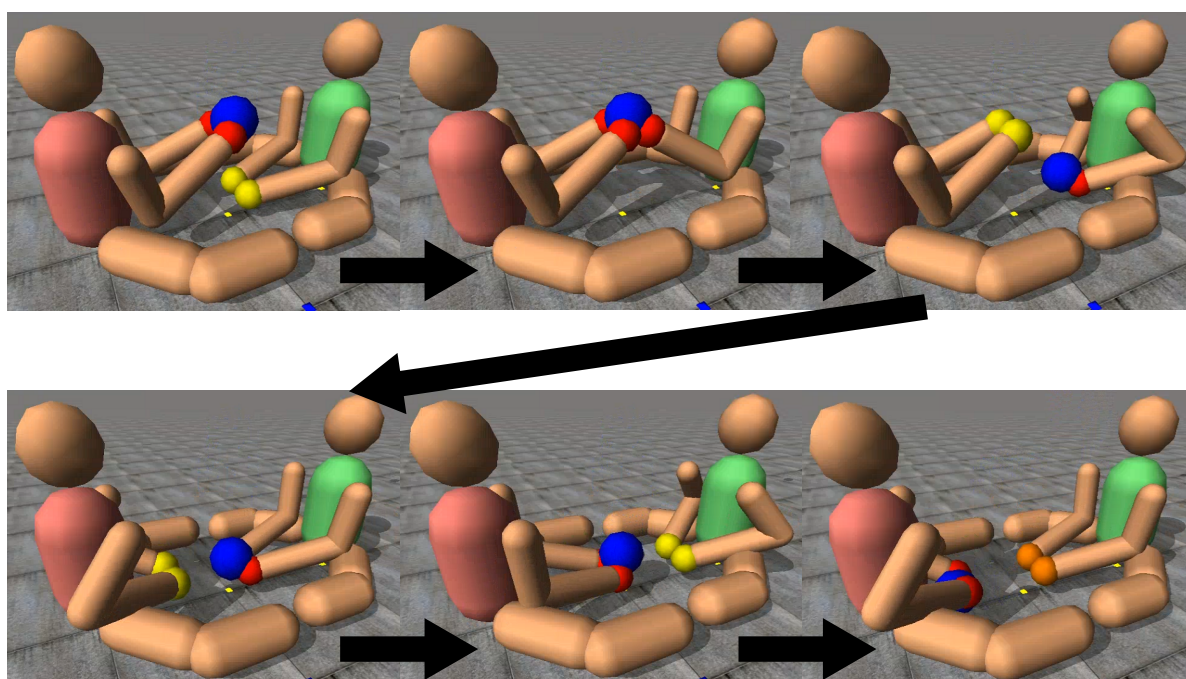


図 3.10: 「ボールの受け渡し」の様子

図3.10はボールをPCが持ち、そのボールをエージェントが受け取り、その後エージェントがPC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．なお、より厳密には最初にエージェントがボールを持ち、そのボールをPCが受け取り、その後PCがエージェントにボールを渡した場合も「ボールの受け渡し」であり、ボールを受け渡す手の位置はFUまたはFLである．また、ボールを渡された側（図3.10においてはエージェント）が1秒以上ボールを保持せずに渡した側（図3.10においてはPC）に返した場合はボールの受け渡しとはみなさないこととした．これはPCおよびエージェントの手先が共にFUまたはFLの位置にあり、瞬間的にボールを握って離すを繰り返した場合を「ボールの受け渡し」とであると認めないためである．

「ボールの受け渡し」は多くの実験参加者で見られた．ボールを使ったやりとり遊びとしても一般的であり、創発された事自体は自然であると言える．しかし、ボールを渡

すだけ、あるいは受け取るだけなどの一方通行のやりとりとは異なり、「ボールの受け渡し」は双方向のやりとりである。双方向のやりとりはやりとりの時間的変化も考慮する必要があるため、提案手法による学習は時間的変化も踏まえたインタラクションの学習ができていることがわかる。

「ボールの押し合い」

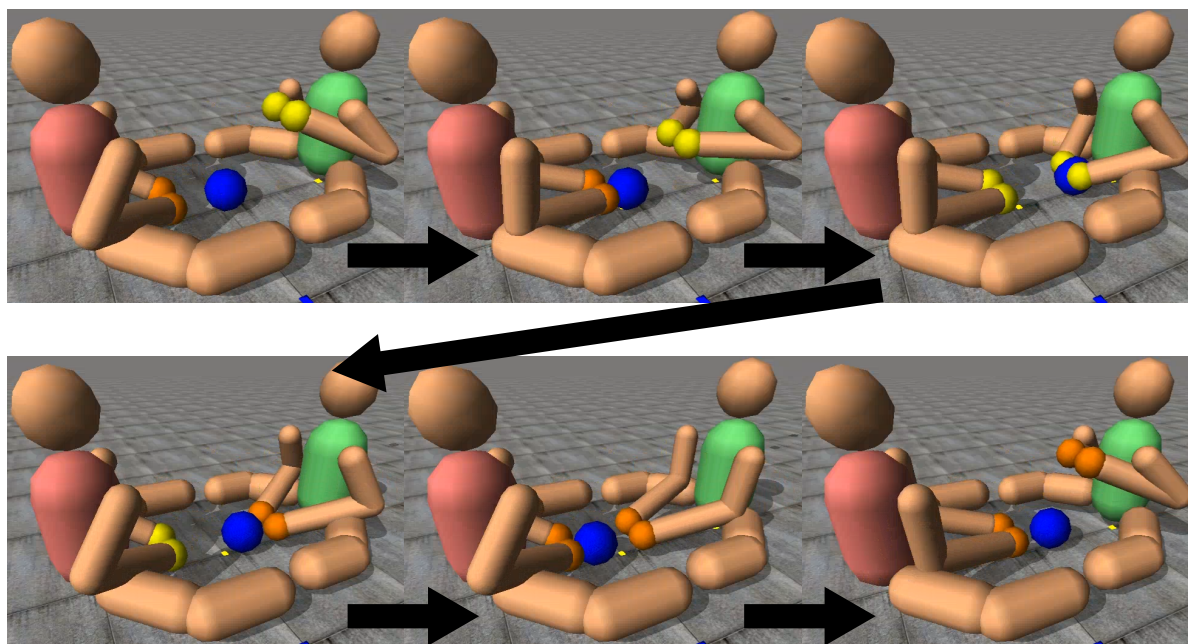


図 3.11: 「ボールの押し合い」の様子

図3.11はボールをPCが握った状態の手でエージェントに向かって押し、ボールはエージェントのもとへ転がり、エージェントが握った状態の手でPCに向かって押しだす様子を示している。これを「ボールの押し合い」と呼ぶ。「ボールの押し合い」はボールを掴まずに相手に渡すインタラクションであり、「ボールの受け渡し」の一種であるとも考えられる。ただし、本環境においては握った状態の手でのみボールを押し出すことができるため、「ボールの受け渡し」とは発生する条件が明確に異なる。具体的にはボールを掴めない位置であらかじめ手を握っておき、握った状態のままボールに接触し相手の下へボールを転がしあうこととする。また、「ボールの押し合い」はボールが地面を転がるのが条件であり、握った状態の手でPCとエージェント間をボールが行き来したとしても、ボールが空中であった場合は下記の「バレーボールのようなやりとり」であるとする。

「バレーボールのようなやりとり」

図3.12はボールをPCが握った状態の手で上方向に飛ばし、それをエージェントが握った状態の手で弾き、PC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

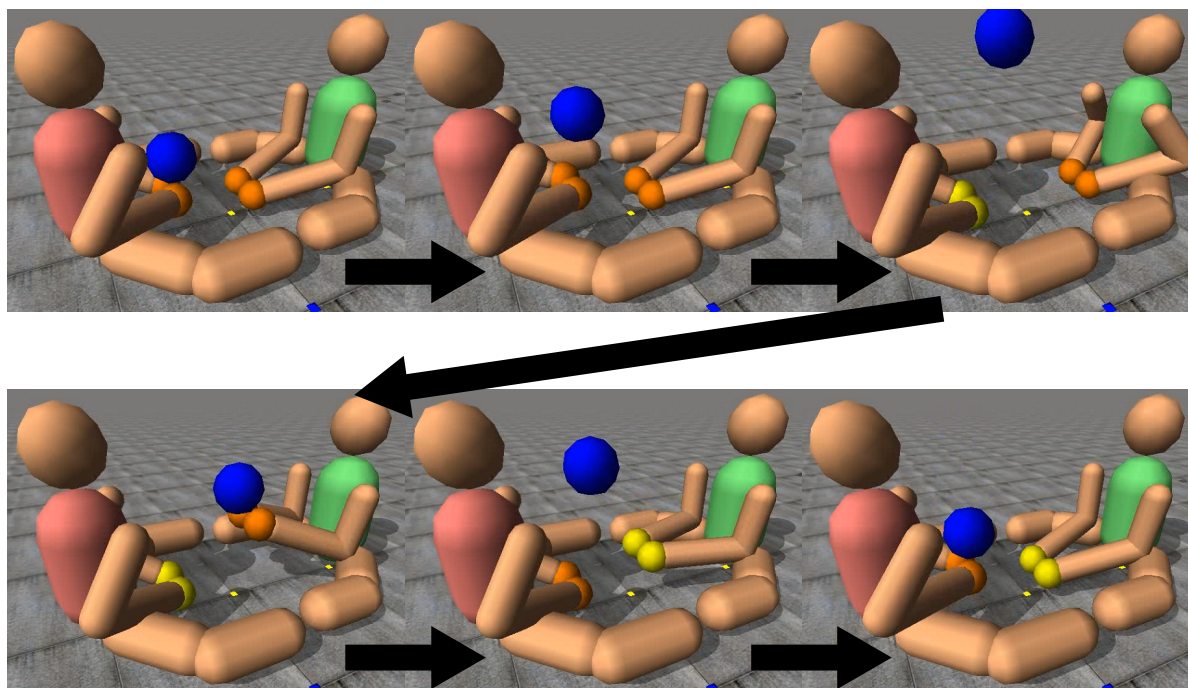


図 3.12: 「バレーボールのようなやりとり」の様子

「バレーボールのようなやりとり」と呼ぶ「バレーボールのようなやりとり」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。「バレーボールのようなやりとり」はボールを使ったやりとり遊びという枠を超え、スポーツやゲームに近いと考えられ、このようなやりとりを創発できることは本章において構築したインタラクション環境は十分に多様なインタラクションを生み出す可能性を持つと考える。特に、ボールを地面に落とさず常に空中にあるよう手で弾くことは、ボールの移動範囲を PC とエージェントの矢状面に固定した本環境においても非常に難しい。そのため、偶然地面に落ちずに PC - エージェント間を往復することはあれど、狙ってそのようなやりとりを学習させることはできないだろうと筆者は実験開始前考えていた。しかし、ある実験参加者は「バレーボールのようなやりとり」を狙って報酬を付与し、実際に何度も「バレーボールのようなやりとり」の創発に成功した。このようなやりとりを学習できたことは、提案手法が本実験環境における多様性に基つき多種多様なインタラクションを学習可能であることを示していると考えられる。

「ボールの引っ張り合い」

図 3.13 はボールを PC とエージェントの双方が掴んでおり、引っ張り合っているやりとりを示している。これを「ボールの引っ張り合い」と呼ぶ。「ボールの引っ張り合い」はやりとりの膠着状態であるともいえるが、人間同士の場合においてはこれもボールを使ったやりとり遊びであると言える。すなわち、「ボールの引っ張り合い」においてはボールそのものではなくボールを引っ張る力をやりとりしている。また、力の大小関係からボールは両者の間を移動するため、全くの膠着状態ではない。しかし、本インタラクシ

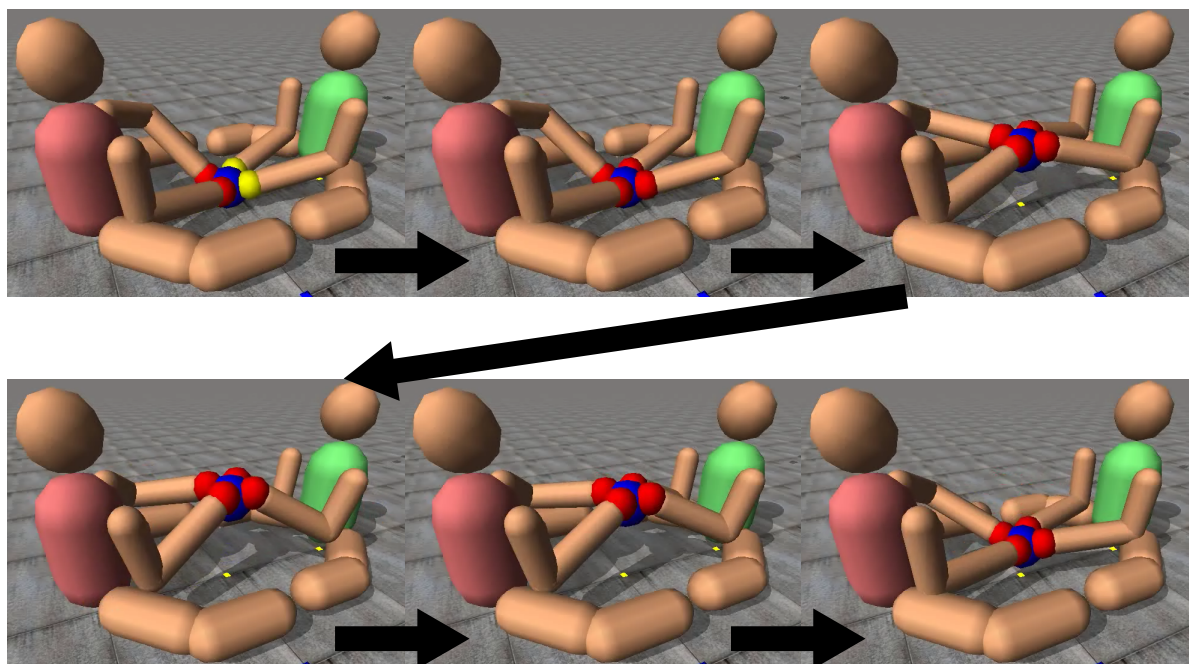


図 3.13: 「ボールの引っ張り合い」の様子

ン環境においてはPCとエージェントの引っ張る力等は同一に設定されており、ボールは左右方向には一切動かない。上下方向には動くため、完全な膠着状態ではないのだが、人間同士の場合と比べて相手のボールを引っ張る力がボールを通して伝わることもないため、やりとり遊びとして捉えることは難しい。「ボールの引っ張り合い」に関しては実機ロボットを用いた場合にやりとり遊びとして有効である可能性が考えられる。

「ボールをぶつける」

図3.14はPCがボールをエージェントに向かって投げ、エージェントの体に当たった様子を示している。逆に、エージェントがPCに向かってボールを投げてPCの体に当たる様子も観測している。これを「ボールをぶつける」と呼ぶ。「ボールをぶつける」ことも双方向のやりとりとして捉えることが出来る。ボールを投げた側はボールに当たる側がボールを掴むことを期待していない。しかし、何度もボールを投げるために自身にボールを返してくれることは期待している。ある実験参加者はエージェントに向かってボールを投げ、体に当たったボールをPCに返すような動きをしたエージェントに対して高い報酬を与えていた。一方、エージェントがPCに向かってボールを投げてしたことに対して高い報酬を与えていた実験参加者もみられた。これはボールを掴んで渡すよりも高度な投げる動作の達成に対して実験参加者が好印象抱いたためと考えられる。エージェントがボールを投げるためには、基本的にエージェントの手先の状態が「NL・握っている(ボールを掴んでいる)」から「FU・握っていない」へ遷移するか、「NU・握っている(ボールを掴んでいる)」から「FU・握っていない」へ遷移する必要があり、非常に限られた状況系列を学習する必要がある。そのため、エージェントがボールを投げるこ

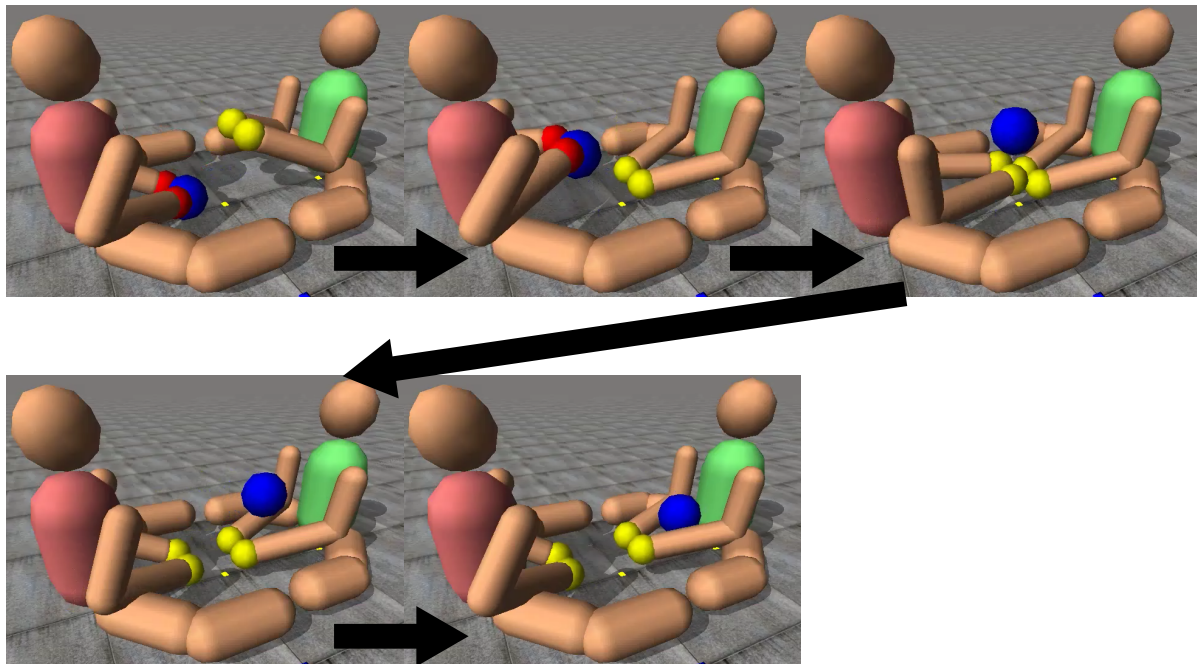


図 3.14: 「ボールをぶつける」様子

とは稀であり，出現時には実験参加者に強い印象を残したと考えられる．

「ボールを防ぐ」

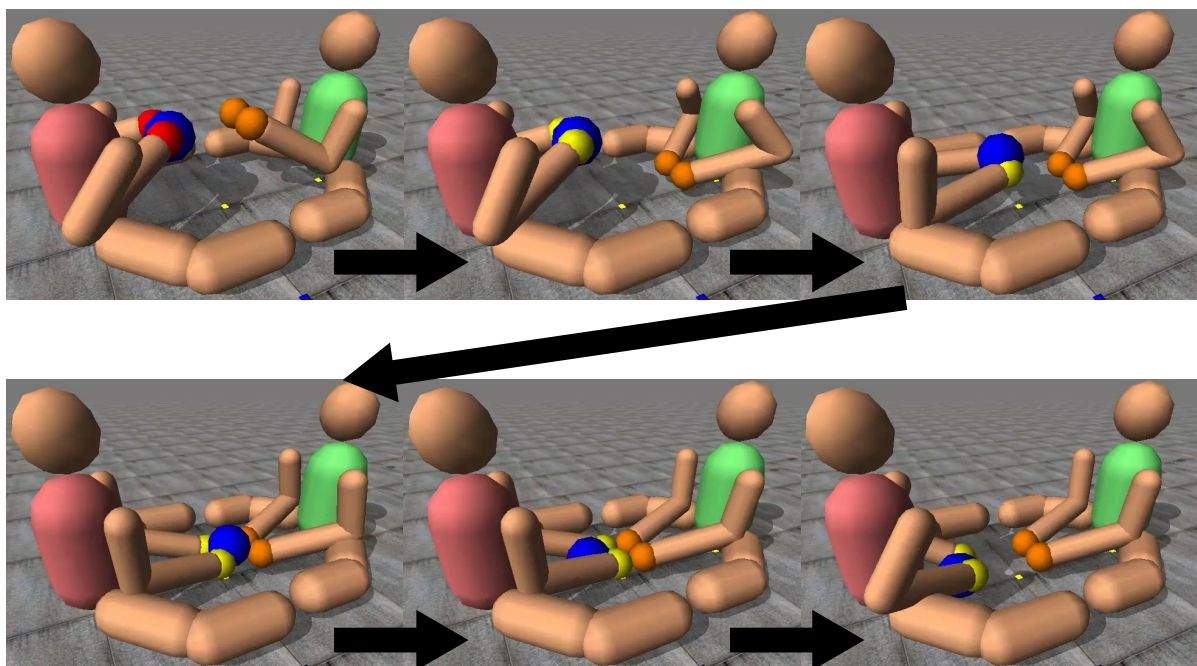


図 3.15: 「ボールを防ぐ」様子

図3.15はPCが投げたボールをエージェントが握った状態の手で弾いている様子を示している。これを「ボールを防ぐ」と呼ぶ。「ボールを防ぐ」やりとりは「ボールをぶつける」やりとりとは違い、エージェントが投げられたボールに反応して握った状態の手で弾く、すなわちエージェント自身の体にボールが当たらないように手先で防ぐことを期待されているやりとりである。一部の実験参加者において「ボールをぶつける」やりとりから「ボールを防ぐ」やりとりに発展している様子を確認した。「ボールを防ぐ」ことが出来れば高い報酬を与えており、特にボールを握った状態の手で弾いた場合よりも、エージェントがボールを掴むことによって防いだ場合に高い報酬を与えていることを確認した。なお、上記の逆パターン、すなわちエージェントが投げたボールをPCが防ぐやりとりに関しては、偶然そのように見える状況の変化だった場合を除き、意図的に学習した様子は見られなかった。エージェント側がボールを投げることは先述の通り非常に稀であり、学習そのものが難しかったと考える。

「エージェントの一人遊び」

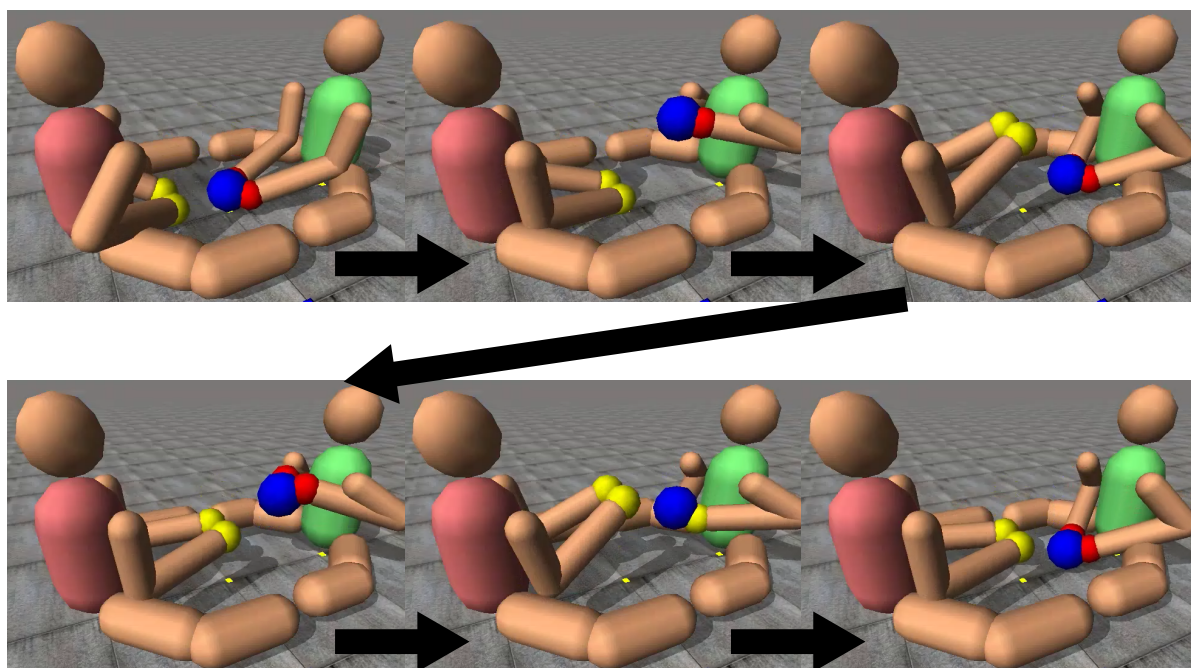


図 3.16: 「エージェントの一人遊び」の様子

図3.16はエージェントがPCの手が届かない範囲でボールを持ち上げたり降ろしたりしている様子を示している。これを「エージェントの一人遊び」と呼ぶ。「エージェントの一人遊び」については双方向のやりとりであるとは言い難い。多くの実験参加者は図3.16のような「エージェントの一人遊び」が見られたら負の報酬を与えていた。しかし、一部の実験参加者はエージェントの一人遊びに対して即座に負の報酬は与えず、じっと様子を観察していた。あるいは、正の報酬を付与する実験参加者も見られた。これは、エージェントを幼児とみなし、幼児の一人遊びを見守る保護者としてPCを操作していたと思われる。エージェント自体の年齢設定および性格などは一切設定していないが、

エージェントとのやりとりから一種のロールプレイへ発展することはメディアイクエーション [14] の観点からも妥当であり、「エージェントの一人遊び」を助長するような報酬付与を行う実験参加者が現れたことも不思議ではない。このように本実験環境は、筆者が想定しなかったやりとりをも創発できる枠組みであり、提案手法は十分にそれを学習出来るものであると考える。

「エージェントのフェイント」

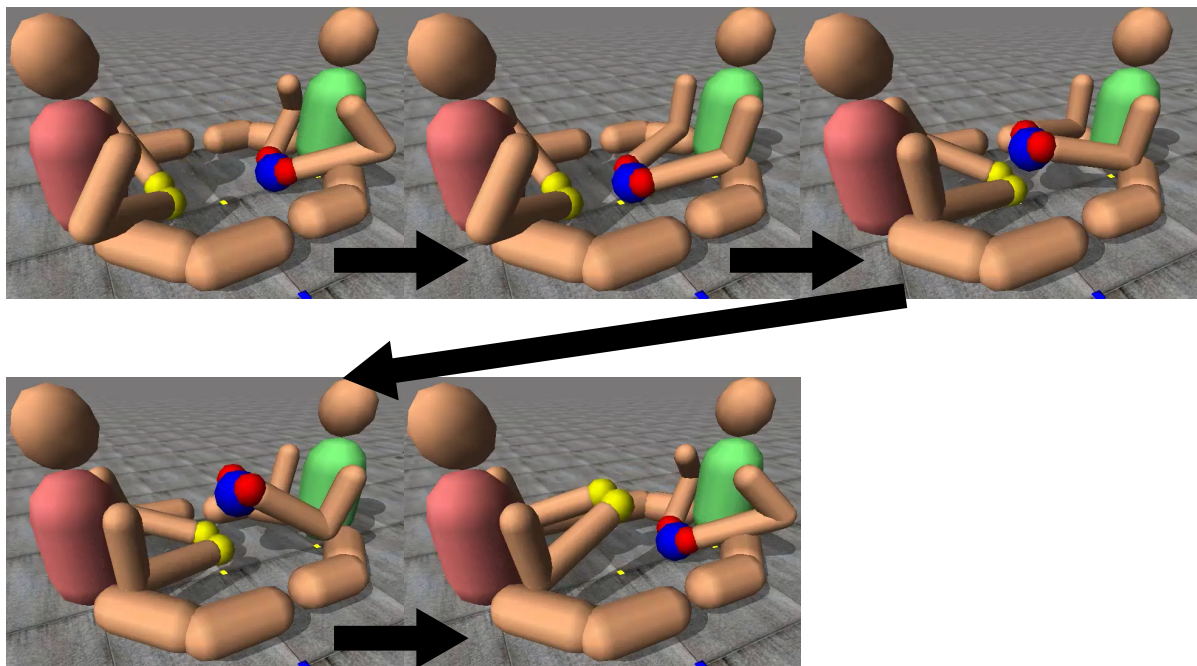


図 3.17: 「エージェントのフェイント」の様子

図 3.17 はエージェントがボールを持った状態で PC に向かって手を伸ばすものの、PC の手から逃げるように手を動かしボールを PC に渡さない様子を示している。これを「エージェントのフェイント」と呼ぶ。「エージェントのフェイント」は「エージェントの一人遊び」と同様、多くの実験参加者が負の報酬を与えるやりとりではあったが、インタラクション相手の存在があって初めて成り立つ双方向のやりとりである。また、「エージェントのフェイント」の逆、すなわち PC がボールを持った状態でエージェントに向かって手を伸ばし、エージェントの手から逃げるように手先を動かす、いわば「PC のフェイント」とも呼べるやりとりも見られた。「PC のフェイント」はそれのみが何度も行われることは無く、例えば「ボールを防ぐ」やりとりを頻出させた実験参加者がボールを投げる前にフェイントをかける、といった形で観測した。このような「エージェントのフェイント」も「エージェントの一人遊び」と同様、本実験環境が多様なやりとりを創発できる枠組みであり、それを学習出来る提案手法の有効性を示すものであると考える。なお、「エージェントのフェイント」に関しては本実験環境構築の際に参考にした栗山ら [40] の実験においても創発が確認されており、全く同一の環境ではないにせよ、同等の創発性を確認できたのではないかと考える。

3.4.5 まとめ

以上の結果から、提案手法が実験参加者の好みを学習し、好むインタラクションを創発し、実験参加者がエージェントに対して好印象を抱く、という本研究の目的達成に成功したと考える。提案手法は「状態と行動の対」を学習する一般的な強化学習と異なり、「状況」というインタラクション相手の行動、自分の行動、環境から得られる観測情報などからなる情報表現を考え、その系列を学習するものである。すなわち、インタラクションの流れ、過去の状況の系列を考慮した行動選択が可能である。これにより、「ボールの受け渡し」や「バレーボールのようなやりとり」といった双方向やりとりが創発され、ひいては実験参加者から「最も印象が良かったエージェント」と評価されるに至ったと考える。

なお、3.4.4 節では提案エージェントにおいて創発されたやりとりを主に紹介したが、「バレーボールのようなやりとり」は近さ優先エージェントでも創発されており、提案エージェントに次いで多く「最も印象が良かったエージェント」と評価された一因ではないかと考える。一方、長さ優先エージェントはあまり良い評価をされなかった。また、ランダムエージェントへの高評価は我々の予想を上回るほどのものであった。これらの印象評価要因を探るため、次節では実験参加者間の報酬付与頻度の違いに着目した。

3.5 実験参加者間の報酬付与頻度の違い

表 3.3: 実験参加者の報酬付与頻度

実験参加者	報酬付与頻度	群
A	0.96	消極
B	2.92	積極
C	0.68	消極
D	1.52	積極
E	1.28	消極
F	0.88	消極
G	1.92	積極
H	1.12	消極
I	2.08	積極
J	0.44	消極
平均	1.38	

3.4.3 節にて述べた通り，本実験では実験参加者は好きなタイミングでエージェントに報酬を与えることができる．そのため実験参加者ごとに報酬を与えるタイミングや回数異なる．表 3.3 に感性評価実験において各実験参加者がエージェントに与えた報酬の頻度を示す．報酬付与頻度が実験参加者平均よりも高い実験参加者は積極的にエージェン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，報酬付与頻度が実験参加者平均以下の実験参加者は「消極群」に分類する．

3.5.1 感性評価の違い

図 3.18 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， t 検定における有意水準 1% および 5% にてエージェント間に有意差を確認できたものを「**」および「*」で，Mann-Whitney の U 検定における有意水準 1% および 5% にて群間の分布に有意差を確認できたものを「‡」および「†」でそれぞれ示す．

図 3.18 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わかる．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有意に上昇していることを確認した（「派手な」: $U = 165$, $p = 0.006$. 「自然な」: $U = 195$, $p = 0.035$ ）．エージェント間の差に着目すると，長さ優先エージェント（エージェント L）の評価に著しい差が確認できる．図 3.19 にエージェント L における積極群および消極群の結果を示す．図 3.19 からわかるように，「派手な」「積極的な」の形容詞対に関しては t 検定により有意差を確認した（「派手な」: $t = 2.484$, $p = 0.038$. 「積極的な」: $t = 3.050$, $p = 0.016$ ）．

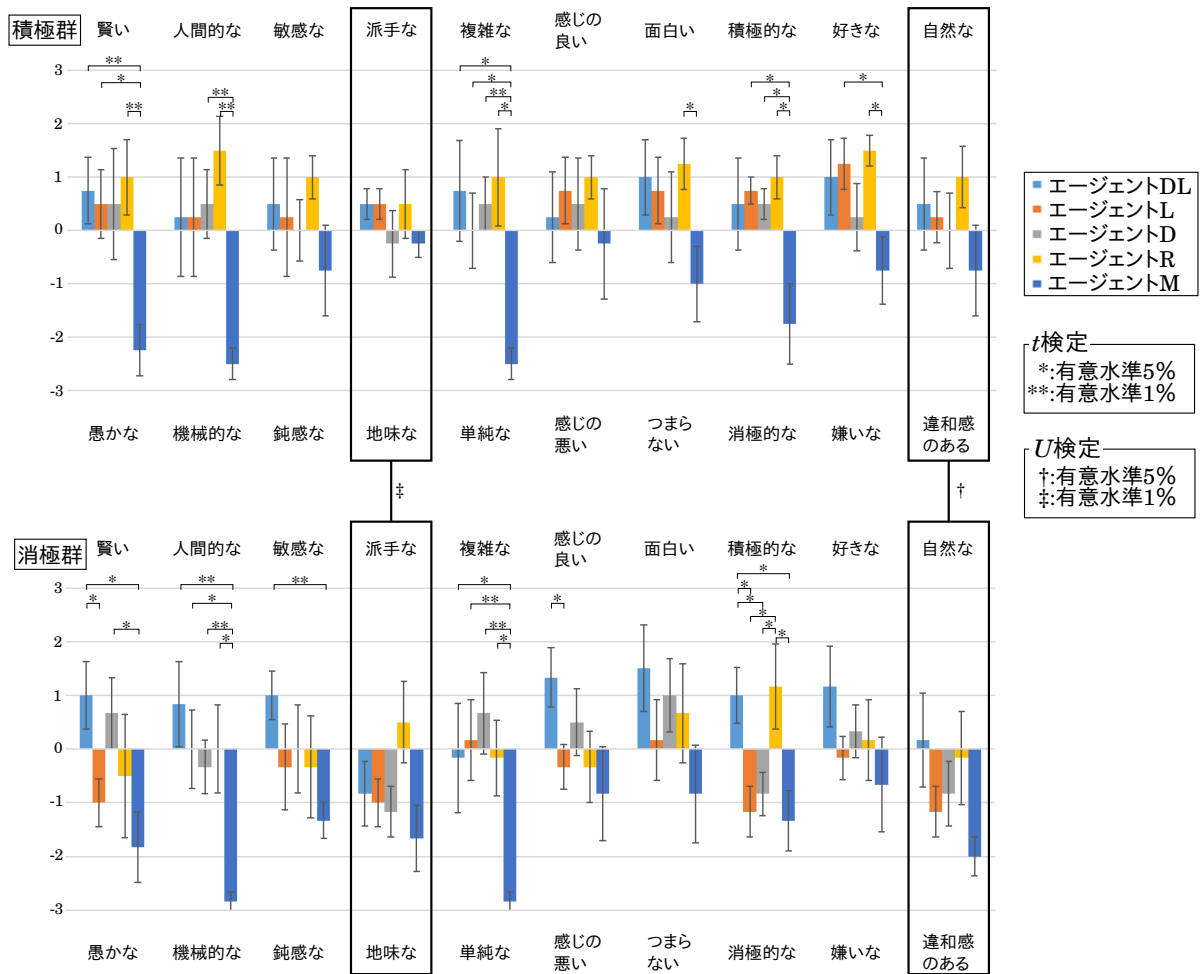


図 3.18: 積極群, 消極群の感性評価実験の結果

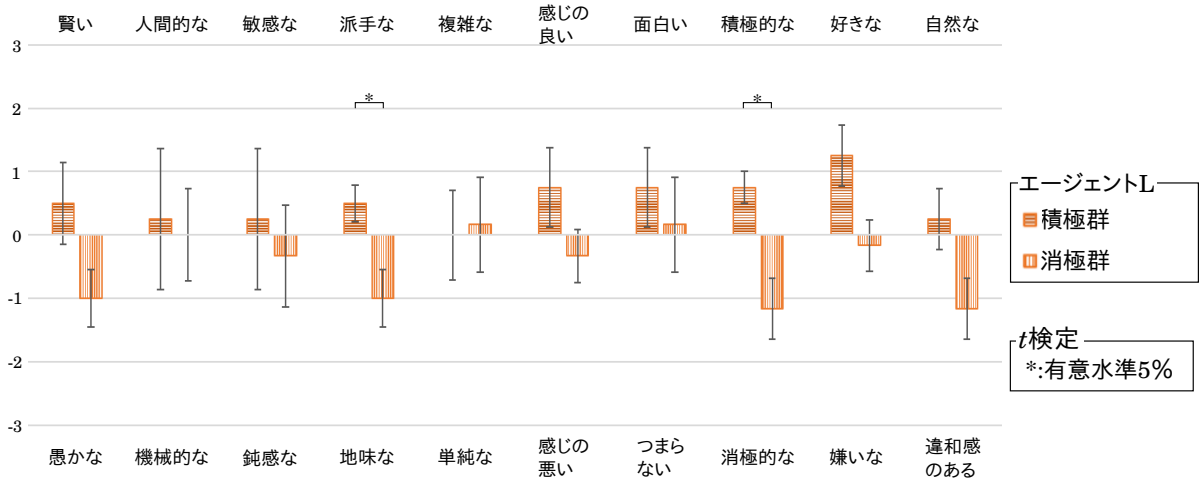


図 3.19: エージェント L における積極群, 消極群の感性評価実験の結果

本実験における全てのエージェントは学習できる系列の長さ制限がある ($T_{max} = 8$). 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、消極群の実験参加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実験参加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 3.1 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 3.1 の影響が大きくなると考えられる。これによりエージェント L に対する群間の差が生じたのではないかと考える。

またランダムエージェント（エージェント R）に着目すると、積極群においてランダムエージェン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られる。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られる。逆に消極群は観測するインタラクションが長く、エージェント 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エージェント R の評価が、積極群と消極群で大きく違うことから示唆される。

3.5.2 インタラクシオンの違い

ここで、各実験参加者群で顕著に見られたインタラクシオンの様子を述べる。積極群はPCが差し出したボールをエージェントが受け取ったら報酬を与える、あるいはエージェントが地面に落ちているボールを拾い、PC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見られたことが、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、成立したとき実験参加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として「ボールの受け渡し(図3.10)」が挙げられる。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っていた実験参加者は10名中6名(実験参加者E, F, G, H, I, J)であり、特に実験参加者Fはミラーリングエージェント以外のエージェント全てにおいてこのやりとりを成立させていた。これは実験参加者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実験参加者は2名(実験参加者G, J)のみであった。

成立しづらい双方向のやりとり例として「バレーボールのようなやりとり(図3.12)」が挙げられる。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」は実験参加者CとJにおいてよく観測された。共に消極群の実験参加者である。特に実験参加者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エージェントへの感性評価は非常にポジティブであった。

3.5.3 まとめ

以上の結果から実験参加者ごとに異なる報酬付与頻度がインタラクシオ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。本実験では各々の実験参加者の性格、例えば他人に物事を教えるようなインタラクシオンにおいて積極的に振る舞うかどうかなどを事前に調査・分類していない。本節における「積極群」「消極群」といった群分けは事後的なものであるため、今後は実験参加者属性を加味した調査を実施したい。また本節および3.4.3節における結果から報酬付与頻度は本研究において重要な要素であることがわかったため、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エージェント、あるいは表情やジェスチャーなどによってエージェントの学習度合いを示すことで、ユーザに報酬付与を促すエージェントを考案していきたい。

3.6 ランダムに振舞うエージェントに対してなぜ好印象を抱くのか

3.4節、および3.5節の結果から着目すべき事象が観測されている。それはランダムエージェントに対する謎の高評価である。ランダムエージェントは2章における二種類の実験ではどちらでも実験参加者から著しくネガティブな評価を受けていた。これは実験の都合上、ユーザの好みを学習するエージェントであると実験参加者に説明した上で実験参加者がエージェントとやりとりを開始するために起こる一種の「適応ギャップ」であり、いわば当然の結果であるとも言えた。なお、適応ギャップとは「ユーザがエージェントに対して期待した機能と実際のインタラクションにおいてユーザが感じた機能の差」と小松ら [51] が定義している。一般的にはエージェントの外見、事前情報からユーザが期待を生成するため、3.4節における実験のように擬人化エージェントを用いて、かつ実験者から「学習する機能を持つ」と事前情報を与えられることによりユーザがエージェントに対して大きな期待を抱くことは容易に想像できる。このことから、本章における実験計画時にランダムエージェントが提案エージェントと同等に良い評価を受けることは予想出来なかった。

本節では人 - エージェント間のインタラクションにおいて、ランダムに行動を決定するエージェントに対して人が抱く印象およびその要因を確認するため、感性評価実験を行い、検証する。

3.6.1 実験設定

実験に用いたインタラクション環境は3.4節と同一である。実験参加者にはいつでも好きなタイミングでエージェントに報酬を与えるよう指示した。報酬値は-5（とても悪い）から+5（とても良い）の間の11段階とし、本実験では「ボールの受け渡し」に関する行動が創発されたと認められる場合ほど高い報酬を与えるよう実験参加者に指示した。「ボールの受け渡し」とは図3.10に示すように、ボールをPC（Player Character）が持ち、そのボールをエージェントが受け取り、その後エージェントがPCにボールを渡す、という一連のやりとりの呼称である。「ボールの受け渡し」はボールを使ったインタラクションにおいて一般的であり、比較的成功しやすいものである。3.4節における実験においても多くの実験参加者で「ボールの受け渡し」が確認されており、実験参加者によっては「ボールの受け渡し」が創発されるようにエージェントに報酬を与えていた者もいた。本実験ではこの「ボールの受け渡し」を目標行動として設定することで、実験参加者間の報酬付与基準を統一した。これは3.4節における実験参加者に見られた「予測できない動きの方が面白い」という印象を排除するためであり、本実験においてエージェントは学習性能を重点的に評価される。実験参加者は20代男性10名。事前にPCの操作方法は説明し、インタラクションに必要な動きを練習させた上で実験に臨んだ。エージェントとのインタラクションは30分間とし、10分毎にエージェントに対する印象を評価させた。

3.6.2 印象評価

実験参加者はエージェントに対する印象をSD法 [45] による感性評価および自由記述により回答する。感性評価は各形容詞対 (図 3.20 参照) について7段階評価で行う。自由記述は「エージェントの学習具合についてどう思いますか」という質問項目に回答させた。

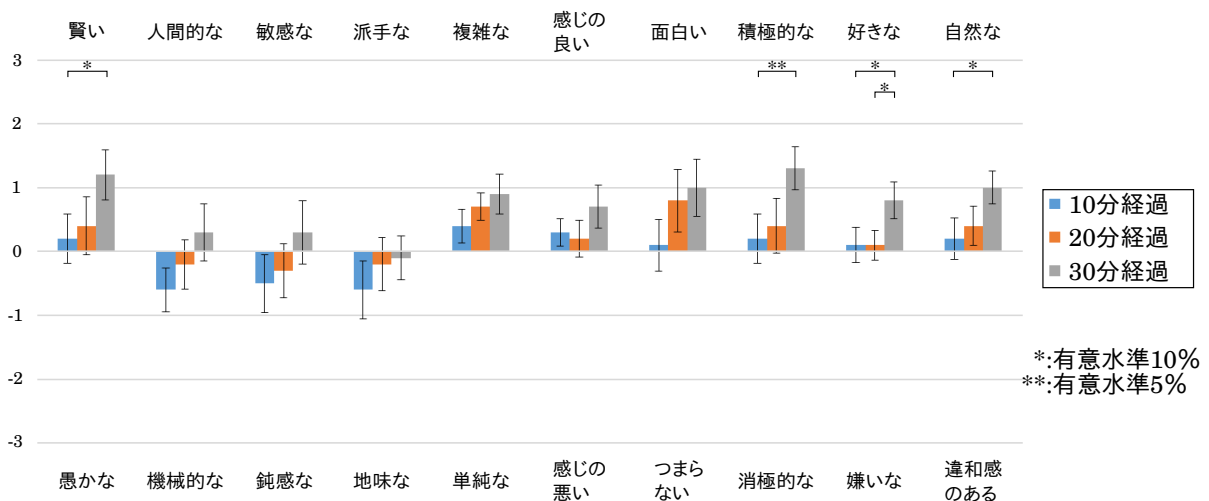


図 3.20: 感性評価実験の結果

表 3.4: 自由記述における内容別回答人数 (実験参加者数 10)

	10分経過	20分経過	30分経過
学習できていない	6名	3名	1名
学習できている	2名	6名	6名
その他	2名	1名	3名

図 3.20 に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す。また各時点の評価に対して t 検定による有意差検定を行った。検定の結果、有意水準 10% および 5% にて有意差を確認できたものをそれぞれ「*」および「**」で示す。

図 3.20 より時間の経過に伴って各形容詞の印象がポジティブに変化していることがわかる。「感じの良い」以外の形容詞対では単調増加であることも確認した。また、10分経過時と比べて30分経過時の感性評価は「賢い」「積極的な」「好きな」「自然な」の4つの形容詞対に関して有意に増加していることを確認した(「賢い」: $t = 1.819, p = 0.086$ 。「積極的な」: $t = 2.144, p = 0.046$ 。「好きな」: $t = 1.744, p = 0.098$ 。「自然な」: $t = 1.922, p = 0.071$)。このことから実験参加者は、与えられた報酬を考慮せずランダムに行動を選択するエージェントにも関わらず、インタラクションによって好印象を抱いたと考えられる。

次に、「エージェントの学習具合についてどう思いますか」の質問項目に対する自由記

述の回答内容に着目し、「学習できている」旨の回答をした実験参加者、「学習できていない」旨の回答をした実験参加者、その他前述の二種類に分別出来ない回答をした実験参加者の3群に分類し人数をカウントした。表3.4に自由記述における内容別回答人数を示す。10分経過時点では「学習できていない」旨の回答は10名中6名であり、「学習できている」旨の回答は2名であったが、30分経過時点では「学習できていない」旨の回答は10名中1名であり、「学習できている」旨の回答は6名であった。このことから実験参加者は、報酬を与え続けることによってエージェントが学習しているような印象を受けており、インタラクション時間の経過に伴い同様な印象を抱く実験参加者が増えたと考えられる。

3.6.3 タスク成功回数と印象

表 3.5: ランダムエージェントにおけるタスク成功回数と自由記述 (実験参加者数 10)

実験参加者	0分～10分		10分～20分		20分～30分	
	成功回数	自由記述の分類	成功回数	自由記述の分類	成功回数	自由記述の分類
A	5回	学習できていない	6回	学習できている	3回	学習できている
B	4回	学習できている	1回	学習できていない	3回	学習できている
D	1回	学習できている	3回	学習できている	0回	学習できている
E	2回	学習できていない	3回	学習できていない	4回	その他
F	2回	その他	5回	その他	4回	その他
C	0回	その他	0回	学習できている	4回	学習できている
G	2回	学習できていない	4回	学習できている	7回	学習できている
H	4回	学習できていない	3回	学習できていない	3回	学習できていない
I	2回	学習できていない	3回	学習できている	2回	学習できている
J	0回	学習できていない	12回	学習できている	3回	その他
平均	2.2回		4回		3.3回	

全平均: 3.17回

表 3.5 に本実験において設定したタスク「ボールの受け渡し」がインタラクション中に成功した回数を 10 分ごとに集計した結果および表 3.4 の実験参加者別内訳を示す。

ここで、与えられた報酬に応じて学習を行い行動を変化させるエージェント、すなわち 3.4 節における提案エージェントに対して同様の実験を行った結果を示す。実験条件はインタラクション相手をランダムエージェントから提案エージェントに変更した以外は 3.4 節と同じである。実験参加者は 20 代男性 10 名。このうち 5 名はランダムエージェントに対する実験にも参加している²。

表 3.6: 提案エージェントにおけるタスク成功回数と自由記述 (実験参加者数 10)

実験参加者	0分～10分		10分～20分		20分～30分	
	成功回数	自由記述の分類	成功回数	自由記述の分類	成功回数	自由記述の分類
A	3回	学習できていない	2回	学習できている	0回	学習できている
B	0回	その他	0回	その他	1回	学習できていない
D	7回	その他	8回	学習できていない	10回	学習できていない
E	1回	学習できていない	3回	学習できている	6回	その他
F	7回	その他	5回	その他	8回	その他
K	7回	学習できている	5回	学習できている	7回	学習できている
L	2回	学習できていない	8回	学習できていない	9回	学習できている
M	0回	その他	2回	学習できている	1回	学習できている
N	17回	その他	10回	学習できている	1回	学習できていない
O	2回	学習できている	1回	学習できている	7回	学習できている
平均	4.6回		4.4回		5回	

全平均: 4.67回

表 3.6 は表 3.5 同様タスク成功回数と自由記述の結果を示したものである。表 3.5, 表 3.6 より、提案エージェントがランダムエージェントに比べタスク成功回数が多いことがわかる。全平均値はランダムエージェント 3.17, 提案エージェント 4.67 であり, t 検定によって 10% 水準で差が有意であることを確認した。しかし自由記述における「学習できている」印象をもった実験参加者の数に着目するとランダムエージェントはのべ 14, 提案エージェントはのべ 13 とほぼ差が無い。すなわちランダムエージェントは提案エージェントと比べ有意にタスク成功回数が少なかったにも関わらず提案エージェントと同等の「学習できている」印象を実験参加者が抱いたと言える。

²実験参加者 A, B, D, E, F の 5 名が両実験に参加している

3.6.4 報酬付与頻度

表 3.7: ランダムエージェントにおける報酬付与回数と頻度 (実験参加者数 10)

実験参加者	0分～10分	10分～20分	20分～30分	平均 (1分あたり)
A	32	32	24	2.93
B	21	25	26	2.4
D	37	31	23	3.03
E	36	36	44	3.87
F	39	39	36	3.8
C	21	23	30	2.47
G	18	30	30	2.6
H	38	52	49	4.63
I	16	8	10	1.13
J	35	32	39	3.53
全被験者平均	29.3	30.8	31.1	3.04

表 3.8: 提案エージェントにおける報酬付与回数と頻度 (実験参加者数 10)

実験参加者	0分～10分	10分～20分	20分～30分	頻度 (1分あたり)
A	15	15	22	1.73
B	26	18	14	1.93
D	35	25	20	2.67
E	46	58	62	5.53
F	41	35	35	3.7
K	24	22	30	2.53
L	23	13	11	1.57
M	17	24	23	2.13
N	46	56	51	5.1
O	7	17	12	1.2
全被験者平均	28	28.3	28	2.81

表 3.7, 表 3.8 に実験参加者が実験中ランダムエージェントまたは提案エージェントに付与した報酬の回数および頻度をそれぞれ示す。表 3.7, 表 3.8 より, 3.4 節における実験時の報酬付与頻度 (表 3.3) と比べて, ランダムエージェント, 提案エージェント共に増加していることがわかる。3.4 節においては実験参加者の主観に基づき「良い」と思うやりとりに対して高い報酬を与えるよう指示したが, 自由に与えられるが故にどのようなタイミングで報酬を与えれば良いのか分からなくなってしまった実験参加者が見られた。一方で, 本節における実験はあらかじめ用意されたタスクを達成したかどうか報酬付与基準であり, 実験参加者にとって報酬付与への抵抗が少なくなったのだろうと考える。

表 3.7 と表 3.8 を比較すると, 提案エージェントと比べてランダムエージェントに対する報酬付与頻度が高いことが分かる。この結果は 3.4 節における実験時の報酬付与頻

度(表3.3)と異なるものであるが、これについては実験時間が影響していると考え、3.4節における実験にて、ユーザの行動を真似するだけのエージェントとのインタラクションにおいて報酬は頻繁に与えられない傾向があることを確認した。報酬付与頻度は報酬付与対象の様子に応じて変化すると考えられる。本稿における報酬はあくまで学習のために与えられるものであり、付与対象が与えられた報酬に鈍感で様子や行動に変化が見られなければ報酬は次第に与えられなくなると考える。3.4節の実験および本節の実験では人型エージェントをインタラクションに用いてはいるが表情や仕草は表出されず、それらからエージェントが学習しているかどうかは読み取れないため、実験参加者はエージェントが学習しているかどうかはエージェントが行った行動のみから推測していると考えられる。つまり、エージェント内部の学習の有無に関わらず行動に明確な変化が見られれば報酬は数多く与えられ、頻度は高くなると考えられることから、本実験で使用したランダムエージェントは実験参加者にとって報酬を与えやすいエージェントであったといえる。一方で、提案エージェントは長時間実験参加者から報酬を与えられることである程度行動選択が決定的に近くなり、次第に表出される行動の変化が少なくなると予想される。これにより、30分間の平均報酬付与回数は提案エージェントと比べてランダムエージェントに対してのものの方が多くなったと考えられる。

3.6.5 ポジティブな印象を抱いた要因

感性評価の結果および自由記述の結果から、実験参加者はランダムに行動を決定するエージェントに対して、報酬によって学習している印象を受け、その印象がポジティブな感性評価に繋がったと考える。実験参加者が本来エージェントは学習できていないにも関わらず「学習できている」という錯覚を起こした要因として考えられるものを挙げる。一つは指示による思い込み、もう一つはピーク・エンドの法則 [52] である。実験参加者は実験者から「ボールの受け渡し」をエージェントが学習できるように報酬を与えるように指示を受けている。必然的に実験参加者はエージェントがなにかしらの方法で与えられた報酬を学習すると思い込み、たとえエージェントの動きがランダムであっても学習した上での動きであると感じてしまうのではないかと考える。同様の現象として、浅田らは実際にはプログラムが確率的に行動を決定するにも関わらず、「人間が行動を決定している」と実験参加者に教示を与えることで実験参加者が人間を想定した思考となることを指摘している [53]。岡田らはエージェントに関する事前刺激がユーザの持つ印象に与える影響を検証し、継続的なインタラクションを通じて事前刺激の影響が強まることを指摘している [54]。すなわち実験参加者の抱く印象は、実際に行われた行動そのものよりも実験者が与えた指示や教示、事前情報に大きな影響を受けてしまう危険性がある。また、本実験環境におけるエージェントが選択できる行動は8種類であり、「ボールの受け渡し」が成功する瞬間も存在する。そのため、偶然「ボールの受け渡し」が成功することにより実験参加者がエージェントは学習しているという印象を強めると考えられる。また、このような偶然の成功はピーク・エンドの法則によって最終的なエージェントへの印象に大きな影響を与えると考える。ピーク・エンドの法則は、自分自身の過去の経験の判定においてピーク時にどうだったか、どう終わったかが重要であるとするヒューリスティクスである。「ボールの受け渡し」の成功をピークとして考えると「ボールの受け渡し」が成功した瞬間の「賢い」、「学習できている」といった印象が最終的な感性評価に繋がっていると考えられる。また印象調査は10分ごとに行われた結果であるため、調査直前のやりとりがどのような印象であったか、すなわちエンド付近の印象が感性評価に大きな影響を与えたと考えられる。

興味深い結果としてランダムエージェントおよび提案エージェントその双方ともにタスク成功回数が0であったにも関わらず「学習できている」印象を抱いた実験参加者がいることが挙げられる。また逆に成功回数が経過時間に伴って上昇しているにも関わらず「学習できていない」印象を抱いた実験参加者もみられる。これらの要因としては Scaffolding が考えられる。Scaffolding はその時々学習者の能力に応じて簡単な学習課題から徐々に難しい学習課題を与えていく方法であり、人が人とのインタラクションを通じて学習する場合に有効であることが知られている [55]。田中らは人がエージェントに教示を与える場面においても Scaffolding が見られること確認しており [56]、本実験においても多くの実験参加者で Scaffolding が見られた。特に見られた場面としては、PC が持っているボールをエージェントが掴んだ瞬間に報酬を与えるといったものである。図 3.21 に PC がボールを持っている状況からエージェントがそのボールを掴んだ状況を示す。

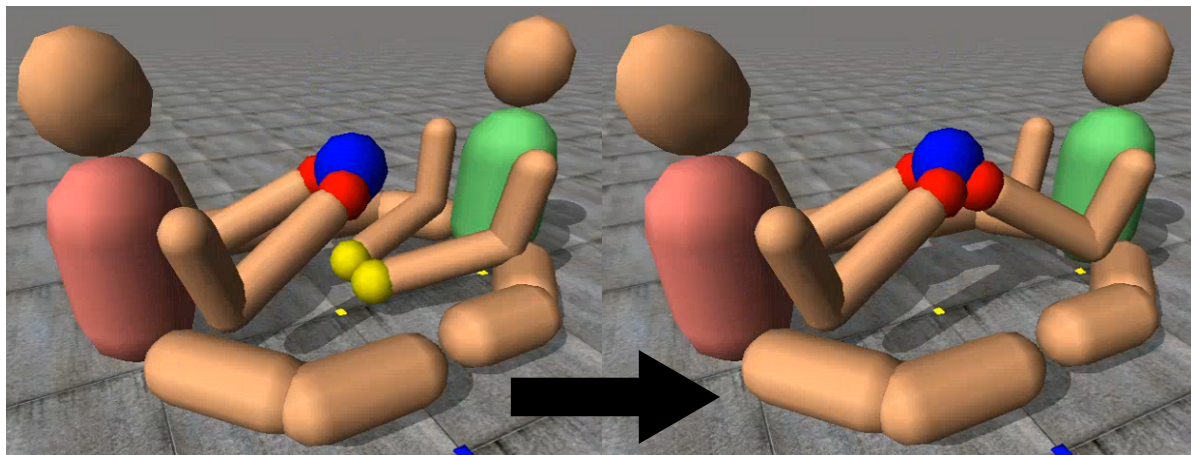


図 3.21: 本実験における Scaffolding 例

報酬付与は「ボールの受け渡し」に関する行動が創発された場合に限る，と実験参加者に指示しており，どのような行動が「ボールの受け渡し」に関するか，と言った点に実験参加者の主観が入る．図 3.21 は「ボールの受け渡し」の中でも最初の場面であり，まずはこの場面を学習させる，すなわち PC がボールを持ってエージェント側に差し出していたらエージェントはそのボールを受け取りに行くという流れを学習させることが重要である．この場面の重要さに気づいた実験参加者はこれだけの状況系列を学習するために何度も高い報酬を与えていた．そして安定してエージェントが PC の持つボールを取るようになってから次の場面の学習へと進んでいた．このような Scaffolding はいわば実験参加者が設定した小さいタスクであり，小さいながらもタスクの達成は実験参加者へ好印象を抱かせる要因となり得ると考える．すなわち，タスクである「ボールの受け渡し」の成功のみがピークとなるわけではなく，実験参加者が独自に設定した別の小さいタスクの成功あるいは失敗がピークとなり，感性評価へ影響を与えたと考えられる．そしてタスクが小さければ小さいほど，行動決定がランダムであることの影響は小さくなるため，ランダムエージェントであっても「学習できている」印象を抱く要因になっていると考える．実際，表 3.5 における実験参加者 C は，10 分～20 分時には多数の Scaffolding と見られる報酬付与を行っており，10 分～20 分時の「ボールの受け渡し」タスク成功回数 0 であるにも関わらず「学習できている」という印象を抱いている．

3.6.6 実験参加者 C の報酬付与の推移と Scaffolding

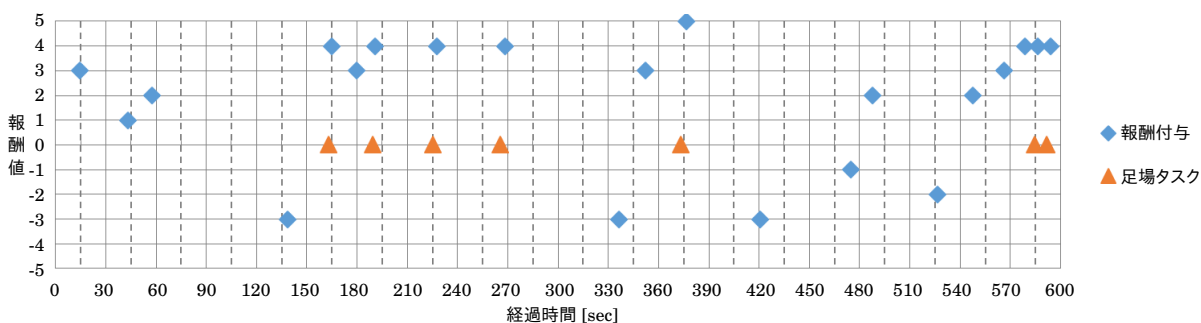


図 3.22: 0 分 ~ 10 分における実験参加者 C の報酬付与および足場タスク発生時間

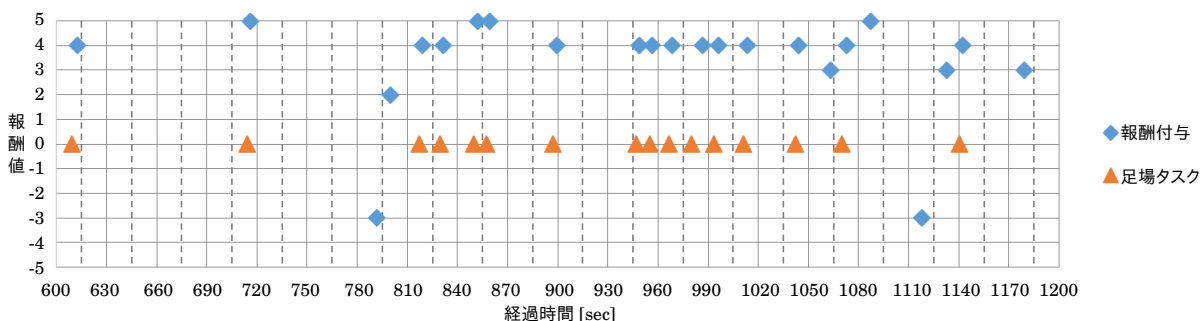


図 3.23: 10 分 ~ 20 分における実験参加者 C の報酬付与および足場タスク発生時間

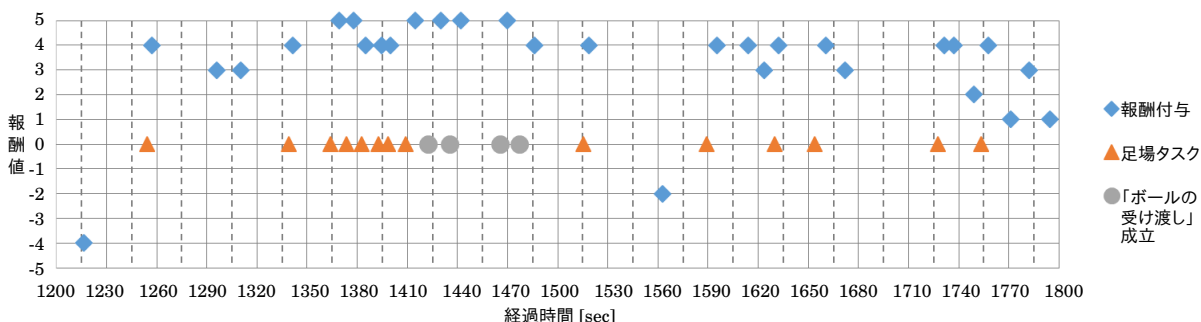


図 3.24: 20 分 ~ 30 分における実験参加者 C の報酬付与, 足場タスクおよび「ボールの受け渡し」発生時間

前節では実験参加者 C を取り上げ, タスク成功回数 0 であるにも関わらず「学習できている」という印象を抱いている原因について多数の Scaffolding と見られる報酬付与の存在を示した. 本節ではより具体的に, 実験参加者 C がどのように報酬付与を行い, 時間経過に伴って Scaffolding と見られる報酬付与がどのように変化したか検証する. 図 3.22, 図 3.23, 図 3.24 に各時間帯における実験参加者 C の報酬付与および Scaffolding と見られる足場タスクが見られた時間を示す. 図 3.22, 図 3.23, 図 3.24 における横軸は実験開始からの経過時間を表している. なお, 10 分, 20 分, 30 分経過時にはそれぞれ実験

参加者に印象調査を行なっているため、各図中の経過時間は実験参加者 C とエージェントがインタラクションを行なっている間の経過時間となっている。また、各図中の縦軸は報酬値を表しているが、「足場タスク」および「ボールの受け渡し」に関しては発生時間を示すことを目的としているため報酬値は 0 としてプロットしてある。実際に「足場タスク」および「ボールの受け渡し」に対して実験参加者 C が付与した報酬値はそれぞれの発生直後にある「報酬付与」プロットを参照していただきたい。「足場タスク」に関しては、本節では実験者が設定したタスクである「ボールの受け渡し」が創発されていないにも関わらず、高い報酬（4 または 5）が付与されたインタラクションとする。ただし、明らかに「ボールの受け渡し」に関する足場タスクではないインタラクションに関しては、高い報酬を付与されたとしても「足場タスク」として扱わない。例として、図 3.22 における 578 秒あたりで実験参加者 C は報酬値 4 と高い報酬を付与しているが、それ以前に行われていたインタラクションは図 3.13 に示した「ボールの引っ張り合い」に近く、「ボールの受け渡し」に関する足場タスクではないと判断した。

では具体的に順を追って見ていく。以下の記述における PC (Player Charactor) とは実験参加者 C が操作する PC のことであり、エージェントとはランダムエージェントのことである。

0 ~ 60 秒：PC はエージェントに向かってボールを差し出している。ボールを掴んだ状態で FU または FL に手先を移動し、エージェントがボールを掴むまで静止する様子が見られる。エージェントがボールを掴んだ時点で PC が手を離し、実験参加者 C はすぐに報酬を付与している（図 3.22 中の 14 秒、43 秒、58 秒）。しかし、エージェントがすぐさまボールから手を離しているから報酬値は低めである。

60 ~ 150 秒：なかなかボールに触ろうとしないエージェントの様子に実験参加者 C は負の報酬を与えている。

150 ~ 270 秒：実験参加者 C が多数の Scaffolding を行なっている様子が確認できる。図 3.25 に図 3.22 中の 163 秒経過時において観測された Scaffolding を示す。図 3.25 で示したように、PC が FL でボールを拾いそのまま FU へ移動、FU で静止していたエージェントにボールを渡し PC がボールを離れた。その直後に実験参加者 C は高い報酬（図 3.22 中の 165 秒では 4）を付与した。このような「ボールの受け渡し」が創発されていないにも関わらず、高い報酬（4 または 5）が付与されたインタラクションを「足場タスク」と呼ぶ。189 秒経過時にも図 3.25 と同様の Scaffolding を観測した。225 秒経過時には図 3.26 に示すような Scaffolding を観測した。これは PC がボールを掴んだ状態で FU で待機し、エージェントが FU に手先を移動しボールを掴んだ所で PC が手を離し、直後に高い報酬を与えた様子を示している。図 3.26 に示したような足場タスクは前節でも述べたように「ボールの受け渡し」の前半部分として重要であり、実験参加者 C 以外にも多くの実験参加者で見られた。266 秒経過時にも図 3.26 と同様の Scaffolding を観測した。

270 ~ 360 秒：PC はボールをエージェントに向かって差し出すように動いている。エージェントの手先に合わせ FU または FL に動かし、エージェントがボールを掴ん

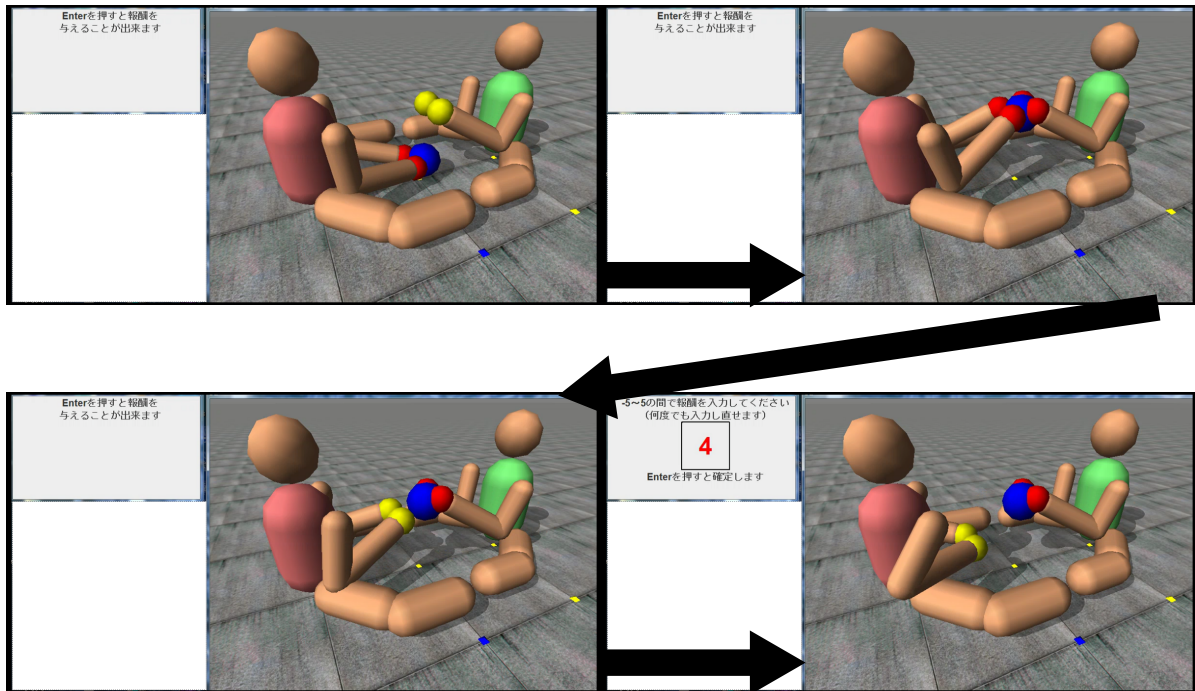


図 3.25: 図 3.22 中の 163 秒経過時において観測された Scaffolding

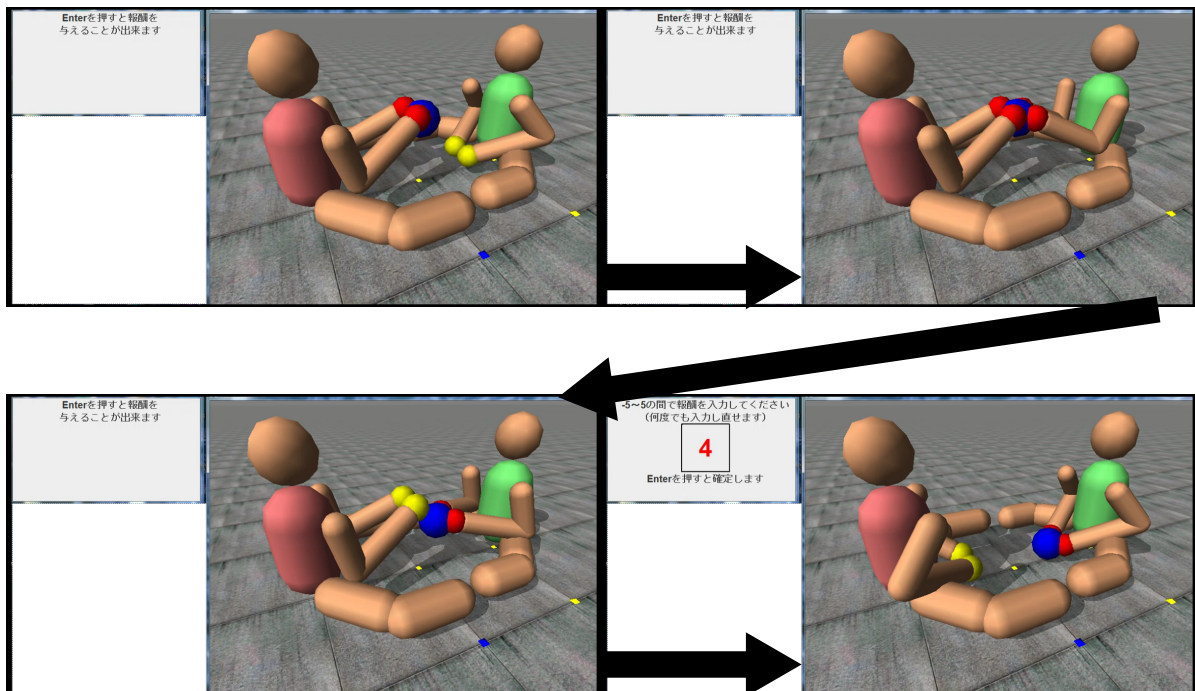


図 3.26: 図 3.22 中の 225 秒経過時において観測された Scaffolding

だらずくに離している．しかしエージェントもまた即座に手を離しているためか、Scaffolding には至っていない．

360 ~ 390 秒：PC がボールを掴んだ状態で NL から FL に動かし，エージェントが FL

でボールを掴んだ後 PC がボールを離した．その直後に高い報酬を付与したため Scaffolding であるとする．FU と FL という高さの違いはあるものの，図 3.26 とおおよそ同じ足場タスクである．

390～570 秒： PC が差し出したボールをエージェントが掴んだ場合においても即座に報酬を付与せずエージェントの動きを待つような様子を観測した．実験参加者 C はこれまでに 5 回ほど Scaffolding と見られる報酬付与をしており，かつ 390 秒までに累計 12 回の報酬付与をしている．すなわち，Scaffolding でいうところの次の足場へエージェントが進むことを期待し，報酬付与を遅らせたと考える．しかし，インタラクション相手のエージェントは一切学習しないランダムエージェントであるため，当然これまでの報酬付与に意味はない．あまりボールに触れようとしない，あるいはボールを掴んだとしても PC に向かって渡そうとしない様子に負の報酬を与えている様子も観測した．

570～600 秒： PC が差し出したボールをエージェントが掴み，ボールを持ったままエージェントが NL へ手先を移動した直後に高い報酬を与えている．これを二回繰り返ししていた．

600～810 秒： PC はボールをエージェントに向かって差し出すように動いているボールを持ったまま FL で静止し，エージェントがボールを取りに来ることを期待しているような動きも観測した．図 3.23 中の 609 秒経過時の足場タスクは図 3.26 と同様であり，714 秒経過時は 373 秒経過時の足場タスクと同様の足場タスクであった．600 秒以前と比べ明らかに報酬付与頻度が下がっているため，エージェントがこれまで与えた報酬を学習できているか観察しているように見える．

810～840 秒： 図 3.26 と同様の足場タスクが見られた．

840～870 秒： 図 3.26 と同様の足場タスクが見られたが，実験参加者 C が付与した報酬値が最高の 5 であった．810～840 秒間との明確な違いはインタラクションの成立の早さが挙げられる．850 秒，857 秒時の足場タスクはこれまでと比べスムーズに成立したように見える．すなわち，余分な状況を挟むことなく遷移したと考えられるため，実験参加者 C は高い報酬を与えたと考える．また，810～840 秒間に与えた報酬を学習し，840～870 秒間においてスムーズに成立したことに對して，エージェントを褒める意味合いで高い報酬を与えたとも考えられる．なお，繰り返しになるがランダムエージェントであるためスムーズに成立したことは偶然である．

870～930 秒： PC はボールをエージェントに向かって差し出すように動いている．エージェントの手先が FL と FU を往復するように動いた際，ボールを持ったまま追いかけるように動く様子を観測した．897 秒経過時の足場タスクは図 3.26 と同様であったが，ボールを FL で掴んだエージェントの手先が FU へと移動し，FL へと戻ろうとする直前に報酬が付与された．これまでの実験参加者 C の報酬付与傾向から，FU または FL で掴んだボールを NU，NL へとエージェントの胴体に近くなるように手先を引く動きをした場合に高い報酬を付与している．FL または FU で掴んだボールを引かずに FU や FL へと移動した場合には実験参加者 C は低い報

酬を付与していた．900 秒時の報酬は 4 と高めであり，実験参加者 C が一度手を引かない場合を許容し始めたのではないかと推測する．

930～975 秒： 図 3.26 と同様の足場タスクを観測した．

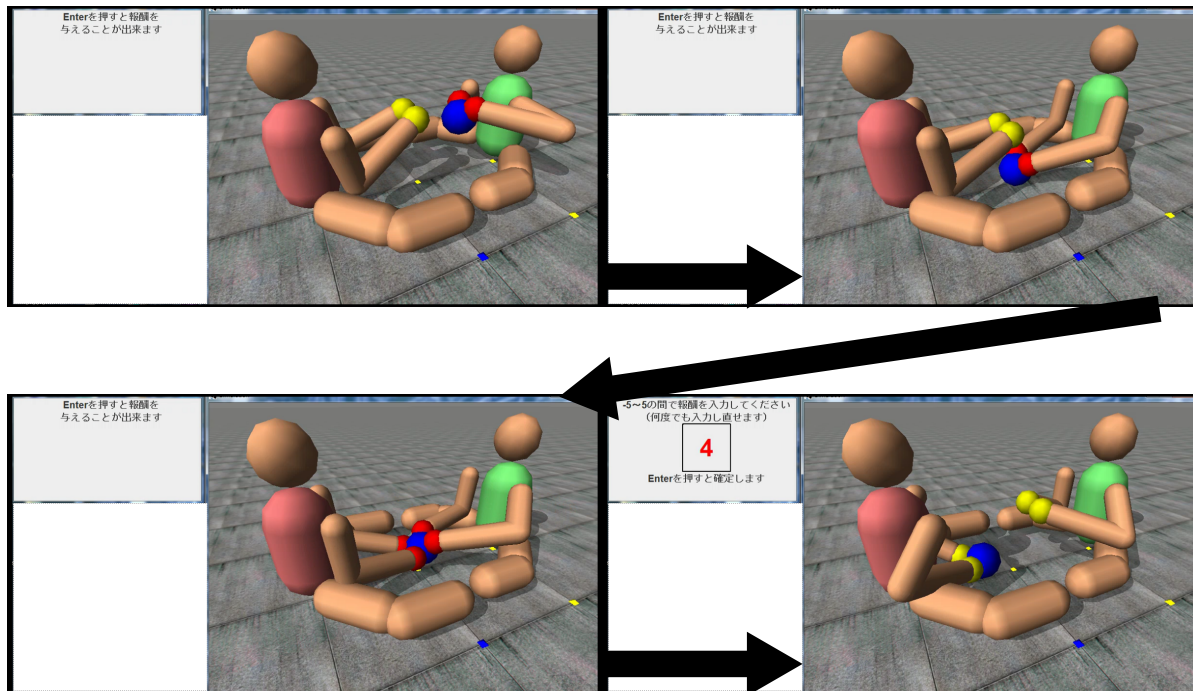


図 3.27: 図 3.23 中の 980 秒経過時において観測された Scaffolding

975～990 秒： 図 3.27 に図 3.23 中の 980 秒経過時において新たに観測された Scaffolding を示す．エージェントがボールを掴んだ状態で NU から FL へと手を移動し，PC がそれを追うように FL へと手を移動．FL から FU へ手を移動しようとしたエージェントを押さえる形でボールを PC が掴み，エージェントがボールを離れた後に高い報酬を付与した．これは「ボールの受け渡し」において後半部分ともいえる受けとったボールをエージェントが返すインタラクションであると言える．最も理想的な形としては PC の手先位置に向かってエージェントがボールを差し出し，それを PC が受け取る状況遷移であるが，図 3.27 に示したようなボールを PC が積極的に取りに行く形でも問題はない．一方，「ボールの受け渡し」をエージェントからスタートしたとも考えられる．どちらの場合でも「ボールの受け渡し」の部分タスクであり，実験参加者が高い報酬を与えた点からも Scaffolding である．

990～1080 秒： 993 秒経過時に図 3.25 と同様の足場タスクを観測した．ただし，ボールは FL で渡された点異なる．1011 秒，1042 秒，1070 秒経過時に図 3.26 と同様の足場タスクを観測した．

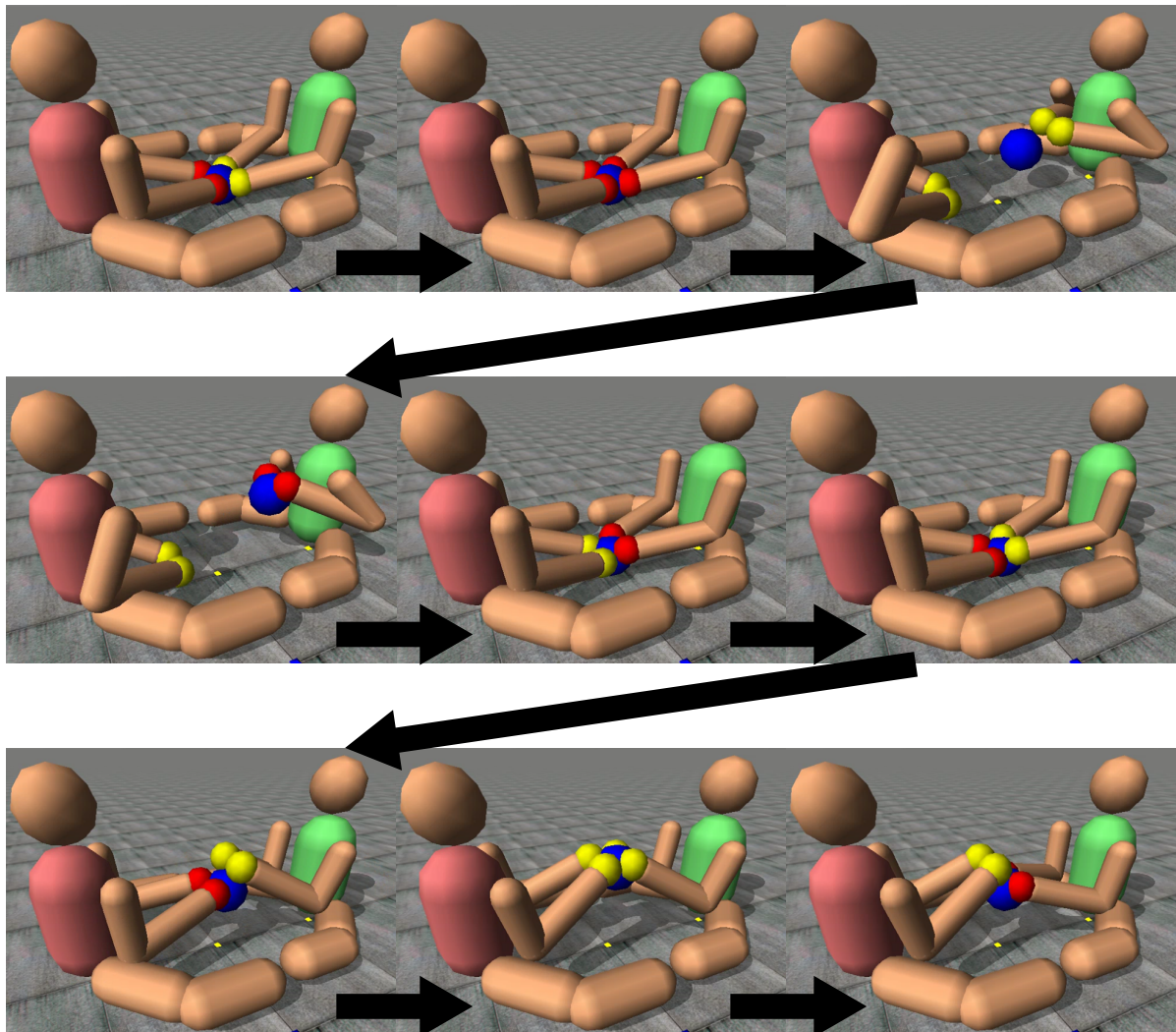


図 3.28: 図 3.23 中の 1081 秒 ~ 1086 秒において観測されたインタラクション

1080 ~ 1095 秒：実験参加者 C は 1087 秒経過時に最も高い報酬 5 を付与している。しかし、その直前に行われたインタラクションは「ボールの受け渡し」でも「足場タスク」でもない。図 3.28 に図 3.23 中の 1081 秒 ~ 1086 秒において観測されたインタラクションを示す。流れとしては、

- (1) PC がボールを掴んだ状態で FL に手先を移動。
- (2) エージェントが FL でボールを掴む。
- (3) その後 PC はボールを離し、エージェントも NU へ手先を移動すると同時にボールを離した。この時ボールには勢いがついたため、エージェントの NU に向かって飛んだ。
- (4) エージェントは NU で手を握った際にボールを掴む。
- (5) 掴んだまま FL へ移動。
- (6) PC も FL へ手先を移動しボールを掴むが、一瞬早くエージェントが手を離した。
- (7) その後、エージェントが FU へ手先を移動し、それを追うようにボールを掴

んだPCが移動する。

(8) PCはFU到着後にボールを離したため、ボールは自由落下を始める。

(9) エージェントがボールを掴む。

上記のインタラクションで重要なのはPCとエージェントの両者がボールを掴んでいる状況が(2)しか無い点である。すなわち(5)から(6)、および(8)から(9)ではボールは「投げ」渡されていることになる。よって図3.28に示したインタラクションは「ボールの受け渡し」には相当しないが(5)から(6)、および(8)から(9)の遷移が一瞬であったために実験参加者Cが「ボールの受け渡し」が成立したと勘違いし高い報酬を付与したと考えられる。

1095～1200秒：PCはボールをエージェントに向かって差し出すように動いている。1140秒経過時に図3.26と同様の足場タスクを観測した。

1200～1350秒：PCボールをエージェントに向かって差し出すように動いている。図3.24中の1254秒経過時、1339秒経過時に図3.26と同様の足場タスクを観測した。

1350～1415秒：1364秒経過時に図3.27と同様の足場タスクを観測した。ただし、エージェントはボールを掴んだままNLからFUへと手を動かし、PCはFLへと移動しようとするエージェントを押さえる形でFUでボールを掴んだ。その後、1374秒、1393秒、1409秒で図3.26と同様の足場タスクを観測した。また、1383秒、1398秒で図3.25と同様の足場タスクを観測した。非常に短い間隔でScaffoldingが見られたが、おそらく1095秒以降エージェントの行動に進歩が見られないことから学習を強めるために行われたと考える。

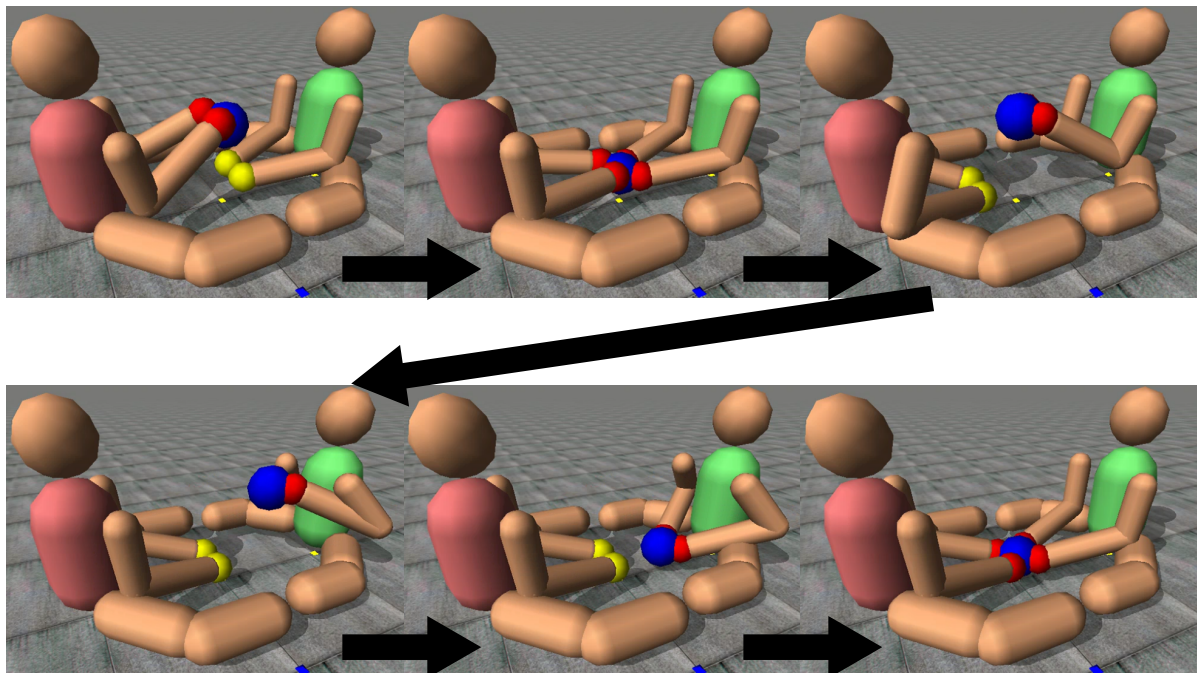


図 3.29: 図 3.24 中の 1422 秒経過時において観測された「ボールの受け渡し」

1415～1455 秒：図 3.29 に図 3.24 中の 1422 秒経過時において観測された「ボールの受け渡し」を示す。PC はボールを掴んだ状態で FU から FL へと手先を移動し、FL にいたエージェントはボールを掴んだ。その後、PC がボールを離し、エージェントはボールを掴んだまま FU へと手先を移動。そのまま NU、NL と移動し、FL へ移動した際に PC もボールを掴んだ。その後エージェントがボールを離したため、「ボールの受け渡し」成立とする。図 3.29 に示した「ボールの受け渡し」は、タスクとして実験参加者に提示した「ボールの受け渡し」(図 3.10) と比べ、エージェントがボールを受け取ってから冗長な状況遷移が含まれている。しかし、エージェントはボールを離しておらず、かつ最短手順のみを「ボールの受け渡し」と設定していないため、間違いなく「ボールの受け渡し」である。タスク成立を受け、実験参加者 C も最高報酬を付与している。1435 秒経過時において観測された「ボールの受け渡し」も、図 3.29 同様ボールを受け取った後のエージェントの行動に冗長さが見られた。

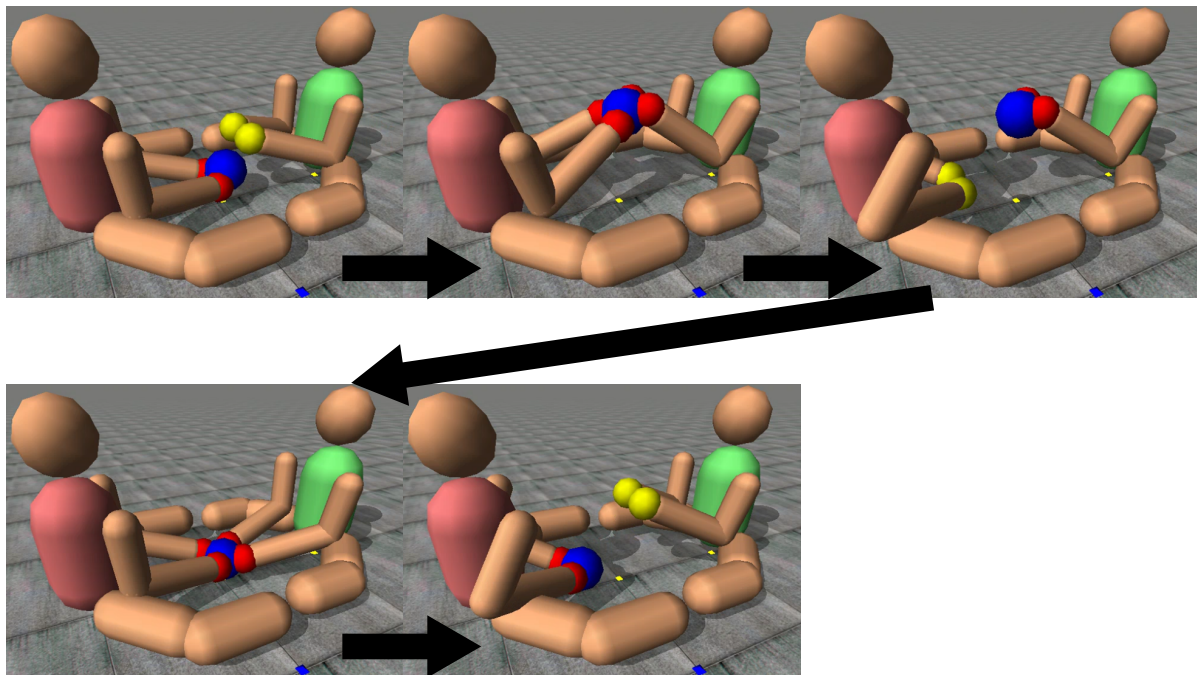


図 3.30: 図 3.24 中の 1466 秒経過時において観測された「ボールの受け渡し」

1455～1500 秒：図 3.30 に図 3.24 中の 1466 秒経過時において観測された「ボールの受け渡し」を示す。PC がボールを掴んだ状態で FL から FU へ手先を移動し、FL にいたエージェントの手先がボールを掴んだ。その後、PC はボールを離し NL へ移動。エージェントは 1 秒ほど FU のままだったが、その後 FL へ移動し、同じく FL へ移動した PC がボールを掴んだ。その後エージェントがボールを離したため「ボールの受け渡し」成立である。図 3.29 とは違い、図 3.30 に示した「ボールの受け渡し」は冗長さがほぼなく、スムーズに成立まで状況遷移したといえる。タスクとして提示した「ボールの受け渡し」(図 3.10) と比べ、エージェントが自身の体側

(NUまたはNL)へ移動しなかったという差異があるものの、「ボールの受け渡し」としての条件は満たしている。また、このようなエージェントの手先が縦移動のみの「ボールの受け渡し」は成立しやすいため、ランダムエージェントであっても多く観測されている。1477秒経過時に観測した「ボールの受け渡し」はエージェントがボールを受けとりPCに向かって差し出すまでは順調に遷移したものの、最後にエージェントがなかなかボールを離さず、「ボールの引っ張り合い」のようなインタラクションも見られた。そのため、実験参加者Cからの報酬が一つ低い4となったと考える。なお、1415秒～1500秒間に連続して「ボールの受け渡し」が成立しているがランダムエージェントであるため偶然である。あるいは、実験参加者Cが本実験環境に十分に慣れた結果、エージェントがボールを離す前にすばやくボールを掴むことで、図3.28に示したような両者がボールをつかんでいない時間が存在することによる「ボールの受け渡し」不成立が無くなったとも考えられる。すなわち、エージェントではなく実験参加者Cの適応が進んでいると考える。

1500～1575秒：実験参加者Cの報酬付与頻度が明確に下がった。これまでと同様にPCはボールをエージェントに向かって差し出しているが、エージェントがボールを掴んでもすぐには報酬を付与せず様子を観察している。これは1415秒～1500秒間に連続して「ボールの受け渡し」が成立したために、エージェントの学習が進行していると実験参加者Cが思い込んでいるためと考える。

1575～1800秒：頻繁に報酬付与を行っており、Scaffoldingも見られる。おそらく実験参加者Cが期待するほどの進歩がエージェントの行動から見られなかったために学習を強めるために行なっていると考える。1589秒、1654秒経過時は図3.27と同様の足場タスクを、1630秒、1727秒、1753秒経過時は図3.26と同様の足場タスクを観測した。

このように、Scaffoldingは数多く見られ、その成立が実験参加者Cに「学習できている」印象を与えたと考える。実験参加者Cに記述してもらったアンケートには、10分経過時には「少しずつ進んでいるように感じる」、20分経過時には「意味のありそうな行動が増えてきた」、30分経過時には「ボールの受け渡しの行動が最終的には成功する回数が増えたように感じた」とあった。また、600～810秒間、1200～1290秒間、および1500～1575秒間は明確に報酬付与頻度が低下し、それまで与えた報酬からのエージェントの学習を期待し、エージェントの行動を観察する様子が見られた。Scaffoldingは段階的に与えるタスクを難しくすることで、より難しいタスクの学習を効率的に行える手法であるが、足場タスクを被学習者がきちんと学習できているか確認してから次の足場タスクを与える必要があるため、上記のようなようすが見られたのだと考える。アンケートはそれまでのインタラクションを改めて思い出してから回答していると考えられ、600～810秒間、1200～1290秒間のようなアンケート回答直後のインタラクションでエージェントの行動を観察するような様子が見られたのは、自身がアンケートで回答した内容の正当性を無意識に確認していた可能性がある。

3.6.7 まとめ

本節ではランダムに振舞うエージェントに対してなぜ好印象を抱くのか、実験を通して検証した。結果として、学習していないランダムエージェントにも関わらず「学習できている」という印象を抱いた実験参加者を確認し、さらにインタラクション時間の増加と共に「学習できている」印象が強化されたことを確認した。これらの要因として実験者による教示、ピークエンドの法則、Scaffolding などの影響が示唆され、ある実験参加者では多くの Scaffolding が「学習できている」印象に強い影響を与えていることが分かった。特に実験参加者 C の報酬付与履歴および付与されたインタラクションを検証することで、タスクとして設定された「ボールの受け渡し」が成立していないにも関わらず「学習できている」という印象を抱いた経緯を確認した。上記影響の排除方法、あるいは効果的な利用方法に関しては今後も検証していきたい。

3.7 おわりに

本章では、ユーザの報酬付与傾向を学習し、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とボールを実装し、提案モデルを持つ擬人化エージェントと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インタラクション実験と感性評価を実施し、提案エージェントが比較対象と比べて良い印象を得たことを確認した。また提案エージェントにおいて、2章における実験と比べ多様なインタラクションが創発できることも確認した。

提案手法は系列のパターンを学習するものである。自己・他者や媒介物などのインタラクション環境を想定し、同環境下で相応しい情報表現を「状況」と設定し、状況の系列を提案手法で学習することで創発されるインタラクションに一般性が生じると考える。具体的には、本実験で想定したインタラクション環境下では、エージェントおよびPCの手の状態、ボールの状態を「状況」として表現することで「ボールを使ったやりとり遊び」を実現した。これ以外にも例えば、案内エージェントの会話制御を想定し、ユーザとの会話履歴を「状況」として表現することで、ユーザが望む会話パターンを学習・生成できると考える。

本実験におけるランダムエージェントに対する印象は想定外に良いものであったが、単に「ランダムだから面白かった」わけではなく、人の心理的特性から「学習できている」印象が増幅された可能性があることが示唆された。特に Scaffolding やピーク・エンドの法則に関しては印象への影響が大きいと考えられるため、エージェントおよびインタラクションの設計時にはこれらを考慮する必要があると考える。特に Scaffolding に関しては、本研究の目標とするユーザの好みを学習するエージェントの設計において避けて通れない問題である。ユーザから与えられた報酬は「足場」として与えられたものなのかどうか。被学習者としてのエージェントがそれらを厳密に判別することで学習効率の向上も考えられる一方、人間側が Scaffolding によって工夫をするのであれば、エージェントは愚直に学習を進めれば良いとも言える。エージェントはユーザから与えられた報酬によってユーザに適応し、ユーザはエージェントの学習具合に応じて報酬付与を調整することでエージェントに適応する。いわゆる「相互適応」を前提としたモデルを構築することで、エージェント設計における学習アルゴリズムの単純化、ひいては設計者の負担軽減に繋がる可能性も模索したい。

第4章

結論

4.1 本稿のまとめ

本稿では、HAIの中でも人間と仮想エージェントのインタラクションに着目し、かつ遊ぶことを目的としたエージェント開発を目指して、ユーザの嗜好を学習しインタラクションに反映する手法を提案した。

2章にてユーザの好むインタラクションを創発するためのユーザ報酬付与傾向を動的に学習する手法として All-Combinatorial N-gram (ACN) を提案した。ACN は N-gram モデルを基としたパターン学習手法であり、その有効性を単純なやりとり遊びを通して検証した。やりとり遊びとして、両者が記号を交互に選び繋げていく遊びに着目し、記号は「短音」を用いたものと「絵」を用いたものの2種類用意し、それぞれにおいて実験を行った。結果として、「短音」「絵」の双方において提案手法を搭載したエージェントは他手法よりも実験参加者から好印象であったことを確認した。また、「短音」のインタラクション実験において、提案手法が実験参加者の嗜好をどのように獲得したか詳細に検証し、個人適応としての有効性を示した。一方で、扱ったインタラクションそのものの単純さから提案手法の応用性、一般性への議論が不十分であったため、ACN およびインタラクションモデルを拡張したものを3章にて提案した。

3章では、より一般化した ACN を提案し、2章におけるインタラクションモデルを拡張したインタラクションモデルにおいて有効性を示した。多様で連続性のある環境としてボールを用いたやりとり遊びを行えるインタラクション環境を構築し、実験を行った。結果として、提案エージェントは他手法を搭載したエージェントと比べ有意に良い印象を実験参加者が抱いたことを確認し、多種多様なインタラクションを創発したことを確認した。基本的な「ボールの受け渡し」だけでなく、筆者が予想しなかった「バレーボールのようなやりとり」をも学習し創発することが出来たことで提案手法の応用性を示せた。また、ランダムに行動を決定するエージェントに対する好印象の要因を追実験にて調査した。これにより、ピークエンドの法則、Scaffolding などの影響が示唆され、今後

のエージェント開発における重要項目であることを示した。特に Scaffolding は、本研究が対象とする「人間がエージェントを教育するインタラクション」において避けては通れない現象であり、Scaffolding の影響を排除するのではなく、積極的に利用することでエージェントの学習コストの軽減、あるいはより高速な学習へと発展させることができる可能性がある。

4.2 今後の課題

今後の課題として、大きく三つ考えられる。

一つ目に「Scaffolding の誘発」。本稿における実験では、Scaffolding に対する配慮が行われていない。Scaffolding とはあくまで指導方法の一つであり、指導対象の理解度に応じて適切な足場タスクを設定する必要がある。すなわち、教師であるユーザがエージェントの理解度を把握する必要があるが、本稿におけるエージェントは全て表情や言葉を表現出来ない。ユーザはエージェントが選出した記号、あるいは手先の移動からエージェントの学習度合いを推し量る必要があった。3.6 節においてランダムエージェントがランダムに行動を選出していることに実験参加者が気づかなかったことはエージェントから得られる情報が少なかったことも挙げられる。よって、エージェントの現在の理解度を表情、インタラクション以外の身体動作（首をひねるなど）によって表現することによって Scaffolding を誘発、あるいは抑制することによってユーザからの報酬付与をより効果的に学習に繋げていきたい。

二つ目に「報酬付与方法」。本エージェントはユーザからの報酬を、ユーザが整数値を入力することにより取得している。これはあくまで GUI 上でインタラクションが完結しているために可能な状態であり、今後扱うインタラクションの発展・多様化に伴い、より抽象的な枠組みでユーザからの報酬を受け取る必要があると考える。例として、ユーザの表情、音声挙げられる。関連研究として宮越らはヒトの表情から感情を判別する手法 [57] を考案しており、Cho らは表情から感情を判別する手法 [58] を考案している。これらの手法のように感情のような数値化できないものを判別する手法はこれまでに多く提案されており、ユーザの好みを表情や音声から推定し、学習することが可能なエージェントは実現可能であると考えられる。このようなエージェントはより実社会に則しているため、有用性が高いと考える。

三つ目に「実験参加者の増加」。本稿における実験は性差の影響を排除するために参加者を全て男性に限定しているが、各実験において性差の影響がどの程度表出するかに関しては検証すべき項目であり、今後は男女混合で十分な実験参加者数を揃えた実験を行いたい、同時に、各実験参加者の特性を調査し、特性毎に検証すべきであると考えられる。特に、3.5 節にて示したように、他人あるいはエージェントに対して教示する場面において、積極的に教示を与える人と与えない人といった違いはその後のインタラクションおよび相手に対して抱く印象に大きな影響を与えられられるため、あらかじめ参加者の特性をビッグファイブモデル [59, 60] 等を用いて調査した上で実験を行いたい。

謝辞

本研究の機会を与え、懇切なる御指導と御鞭撻を賜りました加藤昇平先生に深く感謝いたします。また、本論文に対し、貴重なご助言を頂きました伊藤孝行先生、小田亮先生に深く感謝いたします。

本研究を進めるにあたり多くの助言を頂き、ご協力頂いた加藤昇平研究室の皆様にも深く感謝し、厚く御礼申し上げます。

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）および特別研究員奨励費（15J06660）の助成により行われました。

参考文献

- [1] Masahiro Fujita and Koji Kageyama. An open architecture for robot entertainment. In *Proceedings of the First International Conference on Autonomous Agents*, pp. 435–442. ACM, 1997.
- [2] Masahiro Fujita. On activating human communications with pet-type robot aibo. In *Proceedings of the IEEE*, Vol. 92, pp. 1804–1813. IEEE, 2004.
- [3] F. Tanaka, B. Fortenberry, K. Aisaka, and Javier R. Movellan. Developing dance interaction between qrio and toddlers in a classroom environment: plans for the first steps. In *IEEE International Workshop on Robot and Human Interactive Communication 2005*, pp. 223–228, 2005.
- [4] N. Mitsunaga, T. Miyashita, H. Ishiguro, K. Kogure, and N. Hagita. Robovie-iv: A communication robot interacting with people daily in an office. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5066–5072, 2006.
- [5] Jory Lafaye, David Gouaillier, and Pierre-Brice Wieber. Linear model predictive control of the locomotion of pepper, a humanoid robot with omnidirectional wheels. In *2014 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pp. 336–341. IEEE, 2014.
- [6] Fumihide Tanaka, Kyosuke Isshiki, Fumiki Takahashi, Manabu Uekusa, Rumiko Sei, and Kaname Hayashi. Pepper learns together with children: Development of an educational application. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pp. 270–275. IEEE, 2015.
- [7] 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.
- [8] 堀田大地, 伊藤夏樹, 竹田泰隆, P Ravindra De Silva, 岡田美智男. Pelat: おぼつかなさを有するロボットと人との関わりについて. HAI シンポジウム, pp. 41–47, 2014.
- [9] 神田崇行, 宮下敬宏, 長田拓, 配川有二, 石黒浩. 人口ロボット相互作用における人型ロボットの外見の影響. 日本ロボット学会誌, Vol. 24, No. 4, pp. 497–505, 2006.
- [10] 山田誠二 (監修・著). 人とロボットの < 間 > をデザインする. 東京電機大学出版局, 2007.

- [11] 翠輝久, 河原達也, 正司哲朗, 美濃導彦. 質問応答・情報推薦機能を備えた音声による情報案内システム. *情報処理学会論文誌*, Vol. 48, No. 12, pp. 3602–3611, 2007.
- [12] 小川浩平, 小野哲雄. Itaco : メディア間を移動可能なエージェントによる偏在知の実現. *Human Interface*, Vol. 8, No. 3, pp. 373–380, 2006.
- [13] 佐藤翔吾, 片上大輔, 稲葉通将, 田中貴紘. ドライバーの心的負担を軽減する擬人化エージェントの開発. *HAI シンポジウム*, pp. 308–316, 2014.
- [14] Byron Reeves and Clifford Nass. *The Media Equation*. CSLI Publications and Cambridge university press Cambridge, UK, 1996.
- [15] Masato Hirose and Kenichi Ogawa. Honda humanoid robots development. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 365, No. 1850, pp. 11–19, 2007.
- [16] Junichi Osada, Shinichi Ohnaka, and Miki Sato. The scenario and design process of childcare robot, papero. p. 80, 2006.
- [17] 柴田崇徳. メンタルコミットロボット・パロとロボット・セラピーの展開. *日本ロボット学会誌*, Vol. 24, No. 3, pp. 319–322, 2006.
- [18] 献大西. ホームユースロボット「wakamaru」. *日本機械学会誌*, Vol. 109, No. 1051, pp. 448–449, jun 2006.
- [19] Kazuya Yoshida. Achievements in space robotics. *IEEE Robot. Automat. Mag.*, Vol. 16, No. 4, pp. 20–28, 2009.
- [20] Xichuan Lin and Shuxiang Guo. Development of a spherical underwater robot equipped with multiple vectored water-jet-based thrusters. *Journal of Intelligent & Robotic Systems*, Vol. 67, No. 3-4, pp. 307–321, 2012.
- [21] Koji Ueda, Takao Inoh, Paulo Cesar Debenest, Ryuichi Hodoshima, Eduardo FUMIHIKO Fukushima, Shigeo Hirose, et al. Development of helios ix: An arm-equipped tracked vehicle. *Journal of Robotics and Mechatronics*, Vol. 23, No. 6, pp. 1031–1040, 2011.
- [22] Yuki Horise, Atsushi Nishikawa, Mitsugu Sekimoto, Yu Kitanaka, Norikatsu Miyoshi, Shuji Takiguchi, Yuichiro Doki, Masaki Mori, and Fumio Miyazaki. Development and evaluation of a master-slave robot system for single-incision laparoscopic surgery. *International journal of computer assisted radiology and surgery*, Vol. 7, No. 2, pp. 289–296, 2012.
- [23] 神田崇行, 平野貴幸, イートンダニエル, 石黒浩. 日常生活の場で長期相互作用する人間型対話ロボット 語学教育への適用の試み . *日本ロボット学会誌*, Vol. 22, No. 5, pp. 636–647, 2004.

- [24] Tomohiro Kawaji, Kei Okada, Masayuki Inaba, and Hirochika Inoue. Human robot interaction through integrating visual auditory information with relaxation method. In *Multisensor Fusion and Integration for Intelligent Systems, MFI2003. Proceedings of IEEE International Conference on*, pp. 323–328. IEEE, 2003.
- [25] 鳩康彦, 佐竹聡, 神田崇行, 今井倫太, 萩田紀博, 安西祐一郎. コミュニケーションロボットによる領域参照のための直示的インタラクションのモデル化. *日本ロボット学会誌*, Vol. 29, No. 6, pp. 532–543, 2011.
- [26] Min Kyung Lee, Sara Kiesler, Jodi Forlizzi, Siddhartha Srinivasa, and Paul Rybski. Gracefully mitigating breakdowns in robotic services. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 203–210. IEEE, 2010.
- [27] Timothy Bickmore and Justine Cassell. Relational agents: a model and implementation of building user trust. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 396–403. ACM, 2001.
- [28] Justine Cassell. *Embodied conversational agents*. MIT press, 2000.
- [29] Justine Cassell. Embodied conversational agents: representation and intelligence in user interfaces. *AI magazine*, Vol. 22, No. 4, p. 67, 2001.
- [30] 石塚満. マルチモーダル擬人化エージェントシステム. *システム/制御/情報: システム制御情報学会誌*, Vol. 44, No. 3, pp. 128–135, 2000.
- [31] 石塚満. 生命的エージェントによるインタフェース/メディア. *情報処理*, Vol. 48, No. 3, pp. 257–263, 2007.
- [32] Yukiko I Nakano, Toshihiro Murayama, Masashi Okamoto, Daisuke Kawahara, Qing Li, Sadao Kurohashi, and Toyoaki Nishida. Cards-to-presentation on the web: generating multimedia contents featuring agent animations. *Journal of network and computer applications*, Vol. 29, No. 2, pp. 83–104, 2006.
- [33] Justine Cassell, Tom Stocky, Tim Bickmore, Yang Gao, Yukiko Nakano, Kimiko Ryokai, Dona Tversky, Catherine Vaucelle, and Hannes Vilhjálmsson. Mack: Media lab autonomous conversational kiosk. In *Proc. of Imagina*, Vol. 2, pp. 12–15, 2002.
- [34] 黒田拓也, 山田誠二, 寺田和憲. オンラインショッピングにおける商品推薦エージェントの外見と振る舞いの関係が購買意欲に与える影響. *人工知能学会論文誌*, Vol. 31, No. 2, pp. G–F78, 2016.
- [35] 本武陽一, 福田玄明, 植田一博. 人とエージェント間での内集団関係形成. *人工知能学会論文誌*, Vol. 31, No. 6, pp. AI30–J, 2016.
- [36] 長尾確. *エージェントテクノロジー最前線*. 2000.

- [37] 伊藤京子, 神月匡規, 石井裕剛, 吉川榮和. キャラクターエージェントをアバタとナビゲータとして利用したネットワークコミュニティの実験. 情報処理学会論文誌, Vol. 44, No. 7, pp. 1812–1827, 2003.
- [38] 竹内勇剛. エージェントメディアエイトドインタラクション. 東京電機大学出版局, 2007.
- [39] 馬場翔太郎, 竹田泰隆, P. Ravindra S.De Silva, 岡田美智男. Inamo: 人と一緒に「遊び」を組織するロボットを目指して. HAI シンポジウム, pp. 53–57, 2014.
- [40] 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.
- [41] Nancy Ratner and Jerome Bruner. Games, social exchange and the acquisition of language. *Journal of child language*, Vol. 5, No. 3, pp. 391–401, 1978.
- [42] 佐藤知正, 中田亨. 人と調和するペットロボットのための対人心理作用技術. 人工知能学会誌, Vol. 16, No. 3, pp. 406–411, 2001.
- [43] 徐金安, 荒木健治, 栃内香次. N-gram を用いた帰納的学習によるユーザの行動予測. 電子情報通信学会技術研究報告. HIP, ヒューマン情報処理, Vol. 102, No. 595, pp. 41–48, jan 2003.
- [44] J.A. Xu, T. Itoh, K. Araki, and K. Tochinal. Evaluation of action prediction method using inductive learning with n-gram. In *Signal Processing, 2004. Proceedings. ICSP '04. 2004 7th International Conference on*, Vol. 2, pp. 1605–1609, aug.-4 sept. 2004.
- [45] Charles Egerton Osgood. *The Measurement of Meaning*, Vol. 47. University of Illinois Press, 1957.
- [46] Carol O Eckerman, Claudia C Davis, and Sharon M Didow. Toddlers' emerging ways of achieving social coordinations with a peer. *Child Development*, pp. 440–453, 1989.
- [47] Carol O Eckerman and Mark R Stein. How imitation begets imitation and toddlers' generation of games. *Developmental psychology*, Vol. 26, No. 3, p. 370, 1990.
- [48] Takatsugu Kuriyama and Yasuo Kuniyoshi. Acquisition of human-robot interaction rules via imitation and response observation. In *Proceedings of the 10th International Conference on Simulation of Adaptive Behavior: From Animals to Animats*, pp. 467–476. Springer-Verlag, 2008.
- [49] 神田崇行, 石黒浩, 石田亨. 人間-ロボット間相互作用にかかわる心理学的評価. 日本ロボット学会誌, Vol. 19, No. 3, pp. 362–371, 2001.
- [50] 岩下豊彦. SD 法によるイメージの測定: その理解と実施の手引. 川島書店, 1983.

- [51] 小松孝徳, 山田誠二. 適応ギャップがユーザのエージェントに対する印象変化に与える影響. *人工知能学会論文誌*, Vol. 24, No. 2, pp. 232–240, 2009.
- [52] Daniel Kahneman. Experienced utility and objective happiness: A moment-based approach. *Choices, values, and frames*, pp. 673–692, 2000.
- [53] 浅田麻菜, 伊藤毅志. メタ相手モデルがエージェントとのインタラクション時の思考過程へ与える影響. *HAI シンポジウム* 2014, pp. D-1 (7–pages), 2014.
- [54] 岡田翼, 今井順一. エージェントに関する事前刺激がユーザの印象の変化過程に与える影響. *電子情報通信学会論文誌 A*, Vol. 97, No. 6, pp. 439–442, 2014.
- [55] David Wood, Jerome S Bruner, Gail Ross, et al. The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, Vol. 17, No. 2, pp. 89–100, 1976.
- [56] 田中一晶, 岡夏樹. Scaffolding (足場づくり) を利用した学習系の構築. *人工知能学会全国大会*, pp. 1D2–7(4–pages), 2008.
- [57] Yoshihiro Miyakoshi and Shohei Kato. Facial emotion detection considering partial occlusion of face using bayesian network. pp. 96–101, 2011.
- [58] Jangsik Cho and Shohei Kato. Detecting emotion from voice using selective bayesian pairwise classifiers. pp. 90–95, 2011.
- [59] Lewis R Goldberg. An alternative” description of personality”: the big-five factor structure. *Journal of personality and social psychology*, Vol. 59, No. 6, p. 1216, 1990.
- [60] Ralph L Piedmont, Robert R McCrae, and Paul T Costa. Adjective check list scales and the five-factor model. *Journal of Personality and Social Psychology*, Vol. 60, No. 4, p. 630, 1991.

付録A

発表論文目録

学術雑誌 査読有り

1. 佐久間拓人, 加藤昇平, 「ユーザ報酬付与傾向を反映する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」, 人工知能学会論文誌, Vol.31, No.6, pp.AG-A (11-pages), 2016.
2. 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 C, Vol.134, No.2, pp.303-311, 2014.
3. 佐久間拓人, 宮越喜浩, 加藤昇平, 「All-Combinatorial N-gram を用いたユーザ評価傾向の動的獲得手法」, 人工知能学会論文誌, Vol.28, No.3, pp.273-278, 2013.

国際会議 査読有り

1. Takuto Sakuma, Shohei Kato, “A Ball Game Typed Human-Robot Interaction Based on All-Combinatorial N-grams”, International Symposium on Artificial Life and Robotics (AROB 2015), pp.266-271, Beppu, 2015.
2. Takuto Sakuma, Shohei Kato, “Acquiring Interaction Rule Reflecting User’s Evaluation Tendency”, 5th International Congress of International Association of Societies of Design Research (IASDR 2013), pp.4362-4370, Tokyo, 2013.

国内会議

1. 佐久間拓人, 加藤昇平, 「学習エージェントの外見とリアクションが教示者の抱く印象に与える影響」, 第18回日本感性工学会大会 (JSKE2016), pp.F23 (4-pages), 東京, 2016.
2. 佐久間拓人, 加藤昇平, 「タスク遂行型インタラクションにおけるロボットに対して抱く印象の変化」, HAI シンポジウム 2015, pp.P-23(5-pages), 東京, 2015.
3. 佐久間拓人, 加藤昇平, 「ランダムに振舞うロボットとのインタラクションに対してなぜユーザはポジティブな印象を抱くのか?」, 第17回日本感性工学会大会 (JSKE2015), pp.F43 (4-pages), 東京, 2015.

4. 佐久間拓人, 加藤昇平, 「ロボットとのインタラクションがユーザの持つ印象に与える影響-ランダム性の認知に着目して-」, 第 29 回人工知能学会全国大会 (JSAI 2015), pp.2K4-OS-14a-5 (4-pages), 函館, 2015.
5. 佐久間拓人, 加藤昇平, 「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」, HAI シンポジウム 2014, pp.G-14 (7-pages), 岐阜, 2014.
6. 佐久間 拓人, 加藤 昇平, 「ユーザの好みを学習するロボットとのインタラクションにおける All-Combinatorial N-gram の有効性」, 電子情報通信学会技術研究報告 HCS2014-58, Vol.114, No.189, pp.85-90, 京都, 2014.
7. 佐久間 拓人, 加藤 昇平, 「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-」, 第 28 回 人工知能学会全国大会 (JSAI2014), pp.1E5-OS-23b-6 (4-pages), 愛媛, 2014.
8. 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」, 情報処理学会研究報告, Vol.2014-ICS-175, No.13 (6-pages), 名古屋, 2014.
9. 佐久間 拓人, 加藤 昇平, 「ボールを使ったやりとり遊びにおけるユーザ評価傾向の学習モデル」, 第 15 回日本感性工学会大会 (JSKE2013), pp.D78 (4-pages), 東京, 2013.
10. 佐久間 拓人, 加藤 昇平, 「All-Combinatorial N-gram を用いたビーズパターン遊びにおける応答規則の動的獲得」, 第 27 回人工知能学会全国大会, No.4D1-4 (4-pages), 富山, 2013.
11. 佐久間 拓人, 加藤 昇平, 「ビーズパターン遊びにおけるユーザ評価傾向の動的獲得」, 情報処理学会研究報告, Vol.2013-ICS-171, No.11 (6-pages), 名古屋, 2013.
12. 佐久間拓人, 宮越喜浩, 加藤昇平, 「ユーザ評価傾向を反映したヒューマンインタラクションの創発」, 平成 24 年度電気関係学会東海支部連合大会, pp.D1-4 (1-page), 名古屋, 2012.
13. 佐久間拓人, 加藤昇平, 「インタラクションの N-gram 表現に基づくユーザ評価傾向の獲得」, 第 14 回日本感性工学会大会, pp.A4-03 (4-pages), 東京, 2012.
14. 佐久間拓人, 宮越喜浩, 加藤昇平, 「ユーザ評価傾向の獲得によるヒューマンインタラクションの創発」, 第 26 回人工知能学会全国大会, pp.3L1-R-12-9 (4-pages), 山口, 2012.
15. 佐久間拓人, 宮越喜浩, 加藤昇平, 「ユーザの評価傾向を反映したコミュニケーションルール獲得システム」, 第 74 回情報処理学会全国大会, pp.4ZC-2 (2-pages), 名古屋, 2012.

付録B

受賞目録

1. 佐久間拓人, 平成 27 年度名古屋工業大学学生研究奨励副学長賞, 名古屋工業大学, 2016.3.2.
2. 佐久間拓人, 平成 26 年度名古屋工業大学学生研究奨励副学長賞, 名古屋工業大学, 2015.2.26.
3. 佐久間拓人, 電子情報通信学会東海支部学生研究奨励賞, 電子情報通信学会東海支部, 2014.6.5.
4. 佐久間拓人, 第 26 回人工知能学会全国大会優秀賞, 「ユーザ評価傾向の獲得によるヒューマンインタラクションの創発」, 人工知能学会, 2012.9.28.