

## 2 者の報酬に基づいたツンデレインタラクションのシミュレーション

Simulate of “Tundere” Interaction Based on Reward of Two Agents

西堀遥輝<sup>1</sup> 坂本孝丈<sup>1</sup> 竹内勇剛<sup>1</sup>

Haruki NISHIBORI<sup>1</sup>, Takafumi SAKAMOTO<sup>1</sup>, and Yugo TAKEUCHI<sup>2</sup>

<sup>1</sup> 静岡大学情報学部 〒432-8011 静岡県浜松市中区城北 3 丁目 5-1

<sup>1</sup>Informatics Department, Shizuoka University 3-5-1, Johoku, Hamamatsu-city, naka-ku, Shizuoka, 432-8011 JAPAN

**Abstract:** 敵対的な態度を取り客観的には不快感を与えているにも関わらず関係性が持続されるインタラクションが存在する。これらのインタラクションでは基本的に相手に損をさせるが稀に報酬を与える場合に利得が高く評価されるため関係性が持続すると考えられ、本研究ではこのようなインタラクションをツンデレインタラクションと呼び、2 者のツンデレインタラクションにおいて、どのようにプロセスを経て、インタラクションが継続し得る関係が構築されているのかを検証することを目的とする。具体的な方法として、相手エージェントとのインタラクションにより受け渡される報酬の値がツン状態とデレ状態に変化し、報酬の累計値によりツン状態やデレ状態が移行するエージェントの振る舞いを定義する。そこからツン状態とデレ状態を有するエージェントが相手エージェントとどのようなインタラクションを生成し得るのかについて、シミュレーションを通して検証する。結果として、相手が不快感を与えても自身から遠ざからなかった場合、後に相手に報酬を与えれば利得が高く評価され関係性は持続し、互いの報酬は増加し、ツンデレインタラクションが再現可能であることが示唆された。今後シミュレーションにより導出されたツンデレインタラクションのモデルを人-人、あるいは人-ロボットで再現した際に、同様の結果が得られ互恵的な関係を結べるのか調査していく。

キーワード ツンデレ コミュニケーション欲求 報酬 互恵 シミュレーション モデル化

### 1. はじめに

我々は日々生活する中で、報酬を得ようとする。そして内部状態によって報酬となり得るかどうか決定する[2,13]。例えば、おなかがすいている人にとって、おいしい食べ物は報酬であり、報酬である食べ物を得ようとする。しかしながら、おなかがいっぱいの人にとって食べることは苦痛に繋がることもあり、食べ物は報酬ではなくなる。また、A さんは B さんにとって良いこと（報酬）だと思ってインタラクションを行うが、B さんにとってそれは邪魔でしかなく報酬と成り得ず、A さんと B さんにとっての報酬が異なることがある。つまり報酬であるか否かは、その人の主観に依存する。

また、人は自身の報酬を得るためだけでなく、他者に報酬を与えるためにインタラクションを行うと考えられる。自身が報酬を得る例として、お金や仕事という報酬を受け

取るために交渉というインタラクションを行い、暇を潰すもしくは相手と話したいということ報酬とし得ようとするのであれば雑談というインタラクションを行う。物を買えば他者が喜ぶと思えば自身のお金や時間を消費し、買えば他者に報酬を与える例として挙げられる。また我々が他者とインタラクションを行う際、自身の感情や相手の感情を推定しインタラクションの仕方を変えることがあり、自身と他者の報酬の相対的な差によって変化してくるのではないかと考えられる[3][4]。

本研究では、他者と関わることで自身が報酬となり得る場合や、関わり方によりその評価（報酬の大きさ）が増減する場合のインタラクションに注目する。

先行研究によれば、人には何か贈り物を受け取ったらお返ししなければいけないという返報性がある[1]。つまり報酬を受け取ったら相手に返さなければならない。このようにギブ&テイクの関係では、互いを思いやり、インタラク

ションを行うことで 2 者の報酬が増え、互いが長期的に win-win になるような互恵の関係を築けることがある[15]. また自身の報酬ばかり追い求め、他人のことを考えない (他人にとっての報酬が増えない・減る) インタラクションばかり続けていると、相手とのインタラクションは長続きしない。つまり他者とのインタラクションにおいて、コストと報酬の均衡に左右されながら、個人の報酬増大を目指すものの、他者の報酬と自身の報酬の均衡も重要視する必要があると言える[14].

しかし、DV (ドメスティックバイオレンス) のように、本来報酬に成り得ないであろう暴力が、振るわれる側にとっての報酬と認知されるケースや[5], ゲインロス効果で示されているようにマイナスの印象を与えた後にプラスの印象を与えることで、印象が、プラスの印象を最初に与えるよりより良い好印象を抱かせることが出来る場合がある[6]. これらのインタラクションは当事者以外の客観的観点から見た時、互恵的な関係であるとは言いがたいと考えられるが、当人たちにとっては互恵的な関係となっていると推察される。本研究では、そのように客観的に見れば報酬を受け取れない・損しているにも関わらず、稀に報酬を与える場合に利得が高く評価されるようなインタラクションをツンデレインタラクションと呼ぶ。しかしツンデレインタラクションの存在は確認されているが、どのようなプロセス、人同士の内部状態を経て互恵となっているのか、明らかになっておらず、そのモデル化はされていない。

そこで本研究は、他者とのインタラクション時に受け渡しされる報酬と、他者への働きかけと、他者からの働きかけに対する選好 (内部状態) から接近・回避行動を生成するモデルに基づき[7], ツンデレインタラクションのシミュレーションと、生成されたインタラクションの検証を行う。具体的には、自身と相手の内部状態と、ツン状態・デレ状態のどちらの状態なのかによって報酬の値や行動が決定され、インタラクションによる報酬の合計値により内部状態やツン状態・デレ状態が遷移するエージェントを設計する。そこから、設計したエージェントが生成するインタラクションについてシミュレーションを行う。このモデルにより、「どのタイミング」で「どのような」ツン行為やデレ行為をすれば他者と互恵になれるのか解明されることが期待される。

## 2. 背景

### 2.1. 2 者間インタラクションにおける報酬

本研究において「他者と関わることを報酬とする。しかし「他者と関わる」といっても A が B に対して暴力を振るい関係性が持たたとしても、B にとっては報酬になるとは考えづらい。よって「他者と関わることを報酬に変化するが、増えるケース・減るケース両方があると

考えられる。また先行研究によれば、「私たちは報酬に対してアプローチし、消費 (コスト) を払うことで報酬を得ている」との指摘があり、実際に得た報酬と消費 (コスト) の差分が実際の報酬になるのではないかと考えられる[2]. 報酬と消費に関する式を式 1 に示す。

$$V_A(b_{t_A}) = \text{Reward}(b_{t_A}) - \text{Cost}(b_{t_A}) \cdots \text{(式 1)}$$

$b_t \cdots t$  時に行われる行動

$\text{Reward}(b_{t_A}) \cdots b_t$  によって得た報酬

$\text{Cost}(b_{t_A}) \cdots b_t$  によって消費した報酬

$V_A(b_{t_A}) \cdots A$  が、 $b_t$  によって得られる報酬の総和

では他者に報酬を与えるために行動するのは、消費がかかるのみで、ただ損なのではないかとも考えられるがそうではない。他者から報酬を受け取ったとき、「恩を返さなければいけない」という返報義務感を生じさせることがある[1]. これは人が周囲との関係を良好に保てるように、「親切にされたら親切を返すべきである」という互恵性の社会的規範に基づいてインタラクションが行われるためだとされている[8]. 銀行のシステムというインタラクションで考えたときに、銀行側の所持するお金を消費 (コスト) として利用者に渡す。その後利用者が借りたお金から利益を上げ、銀行に返済する際には貸した金額 + 利子が追加され、インタラクションを行う前と後の、銀行と利用者の報酬量は増加している。つまり、他者に報酬を与えるということは、その時だけ見れば、自身は報酬を得ない・減らしてしまう可能性があっても、後に他者からも自身に対して報酬のフィードバックがあり、長期的な互恵の関係が結べるのではないかと考えられる。対して、相手の報酬を考えずに、自身の報酬ばかり得られるインタラクションをしていると、そのインタラクションによって相手は報酬を得られず、相手がそのインタラクションをするメリットはなく、自身が相手にインタラクションを望んでも拒否される可能性が考えられる。

また、報酬の受け渡しによるインタラクションは必ずしもゼロサムではないと考えられる。特にインタラクションを行うことによって得られる精神的な報酬である、「快」は、お金のように物理的なものではなく、その人の受け取り方によって報酬量が増えることもある[16].

そのような報酬量の変化するインタラクションとして、DV は本来マイナスの報酬であるものをプラスであると認知することがある。DV とは一方 (加害者) がもう一方の配偶者 (被害者) に対して暴力などを振るうが、被害者側が加害者に依存し、離れることが出来ない 2 者の心理的状态を示す[5,9]. ハネムーン期と呼ばれ、加害者側が暴力を振った後、一転して被害者に過剰なほど優しく接する状態が原因だと考えられる。被害者側はこのハネムーン期と暴力を振るわれる時期の、アメとムチの繰り返しにより、

「暴力であろうと相手がかかわってくれるだけでよい」という判断になり、共依存の関係が出来ると考えられている。

またゲインロス効果によれば、マイナスの印象を与えた後にプラスの印象を与えることで、印象が、プラスの印象を最初に与えるよりより良い好印象を抱かせることが出来る。逆にプラスの印象を与えた後にマイナスの印象を与えることで、印象が、マイナスの印象を最初に与えるよりより悪い印象を抱かせることが出来る [6]。

つまり自身が相手にとっての報酬だと思って渡したものが実は相手にとって、莫大な報酬として受け止められることや、むしろ損として（相手の元々持っている報酬量が減る）受け止められることもあると考えられる。

式2にAから行動と報酬を受ける側Bの報酬に関する式を示す。表1にゲインロス効果によるexpの値の変化の表を示す。

$$V_B(b_{t_A}) = f(b_{t_A}) * exp \dots (式2)$$

$f(b_{t_A}) \dots$  Aの振る舞いによってBに渡される報酬  
 $exp \dots f(b_{t_A})$ に対するゲインロス効果。  
 $V_B(b_{t_A}) \dots$  BがAの行動によって増減させられる報酬

表1:t時におけるゲインロス効果による報酬の変化

Bにとって		t時のBにとっての報酬量の変化
(t-1)時	t時	
+	+	+
+	-	--
-	+	++
-	-	-

## 2.2. 報酬とコミュニケーション欲求

本研究では、先行研究のモデル[7]を基に、コミュニケーション欲求を定義する。コミュニケーション欲求は、Acceptance（自身に対する欲求、誰かに関わりたい）とControl（他者に対する欲求、誰かと関わりたい）の2つの独立した変数からなる内的状態であるとし、それぞれの変数の程度により、相手に対するスタンスが決定するものであると定義する。

- 互恵的とは、Control,Acceptance=（高，高）より、自分から働きかけたいし、相手からも働きかけられたい状態である。また互恵とは、互いに恩恵などを与え合うことであり、自身の報酬と他者の報酬が増えるように行動すると考えられる。
- 支配的とは、Control,Acceptance=（高，低）より、自分から働きかけたいが、相手から働きかけられたい状態である。また支配とは、他者を自己の意思、命令、行動に服従させることであり、自身の報酬を増やし、相手の報酬を減らすように行動すると考え

られる。

- 服従的とは、Control,Acceptance=（低，高）より、自分から働きかけたくないが、相手から働きかけられたい状態である。また服従とは、他者からの意思、命令、行動に従うことであり、自己の報酬を減らし、相手の報酬を増やすように行動すると考えられる。
- 接触忌避的とは、Control,Acceptance=（低，低）より、自分から働きかけたくないし、相手からも働きかけられたくない状態である。また忌避とは、他者を避けることであり、インタラクションを行おうとせず、少なくとも自身と他者の報酬を増やさない行動をすると考えられる。

これにより、他者とのコミュニケーションに対する態度としての内的状態を数値的に表現できるようになると考えられる。その図を図1に示す。また報酬とコミュニケーション欲求の関係表を表2に示す。

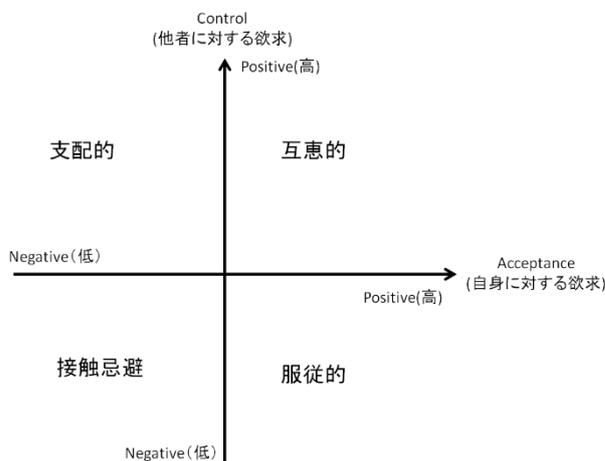


図1:コミュニケーション欲求の2次元での表現

表2:報酬とコミュニケーション欲求の関係表

	Control	Acceptance	自身への報酬	他者への報酬
互恵	高	高	+	+
支配	高	低	+	-
服従	低	高	-	+
忌避	低	低	-	-

## 2.3. 報酬とインタラクションの変化

人が自分や他者の心の動きを類推し、他者が自分とは違う信念を持っていることを理解する機能のことを「心の理論」と呼び、人は他者とコミュニケーションを取るとき、この心の理論を用いて相手の心の内部構造を推論し、自分の行動を決定していると言われている[1]。また他者とコミュニケーションを取る際、他者との衝突をうまく避けることができ、適応的に行動が可能であると言われている[10,11]。つまり他者とのインタラクションにおいて、他者

の報酬量を推定し、それに基づいてインタラクションの取り方を変えることが出来るのではないかと考えられる[12]. しかし自身の報酬量も増やしたいという観点から[1], 自身の報酬量もインタラクションの変化の要因であると考えられる. つまり, インタラクションのとり方(コミュニケーション欲求)の変化は, 自身と他者の報酬両方が要因として存在し, 2つの報酬の相対的な差によって変化してくるのではないかと考えられる.

## 2.4. ツンデレインタラクション

### 2.4.1. ツンとデレの定義

ツンデレインタラクションは, ツン状態とデレ状態という二つのモードに該当する行動をすることで行われる. そこでツンデレ行為を行う側を A (以降 A), 行為をされる側を B (以降 B) とし, A から B に対する行動を Action と定義する. 例えば, Action A+ B+であれば, A と B にとっての報酬が同量に+となるような A から B に対する行動をするという意味になる.

- デレ状態

デレ状態とは他者に対してデレた態度, つまり相手に甘える・話しかけて欲しそうにする・寄りそう・優しくする状態である. 相手に優しくするためには, 相手の報酬量が増えるようなインタラクションを行えばいいと考えられる. よって該当するのは以下の Action だと考えられる (表 3).

表 3 : デレ状態の Action

状態変化	Action A→B	報酬量変化	
		A	B
デレ状態→	A- B+	A-	B+
デレ状態	A+ B++	A+	B++
ツン状態→	A- B+	A-	B++
デレ状態	A+ B++	A+	B+++

- ツン状態

ツン状態とは, 他者に対してツンとした態度, つまり相手を自身から遠ざけようとする状態である. 相手を遠ざけるためには, 相手の報酬量を減らす or 相対的に見て相手の報酬量が増えていないと感じるようなインタラクションを行えばいいと考えられる. またデレ状態からツン状態に移行した際の Action では, ゲインロス効果により B の受け取る報酬量は減少すると考えられる. またツン状態とデレ状態を何度も往復することで, B にとって A から関わられること自体が報酬となり, 本来 B の報酬量が減るようなツン状態でも, B の認知する報酬量が増加するケース

も考えられる (表 4).

表 4 : ツン状態の Action

状態変化	Action A→B	報酬量変化	
		A	B
ツン状態→	A+ B-	A+	B-
ツン状態	A++ B+	A++	B+
デレ状態→	A+ B-	A+	B--
ツン状態	A++ B+	A++	0
デレ→ツン	A+ B-	A+	B+
(往復)	A++ B+	A++	B+

### 2.4.2. ツンとデレの状態遷移

ツン状態は常に A が報酬を得ており B は常に報酬が減っており, A と B との報酬量の差は広がり続ける. しかし A はある程度自身の報酬が貯まるとデレに遷移するのではないかと考えられる. 今回のモデル化においては, A の報酬量が B の 2 倍になっていれば, A は十分満足し, B は不満足となるため, デレ状態に遷移するものとする.

デレ状態は常に B の報酬が, A と相対的に見たときに増えていく. これも A がある程度, B が満足していると予測したところでデレからツンに遷移し, 今回のモデル化においては, A と B の報酬量がほぼ同じ量になればツン状態に遷移するものとする.

また, B の報酬量が減るようなツン状態でも, ツンとデレ状態を交互に繰り返すことによって, B の認知する報酬量は増加する状態も存在する. この状態は, ツン状態とデレ状態を往復した回数に応じて移行する可能性が上がるものとする (ツン 2). またこの状態から本来の報酬が減ると認知するツン状態に移行することも考えなければならない. これもツン状態のループ回数に応じて遷移しやすくなるものとする (ツン 3).

ポライトネス理論に関する研究において, インタラクションを行う 2 者にはそれぞれ対人欲求を示す Face が存在し, FTA(Face-threatening act)と呼ばれる Face を侵害する行為をされた場合, インタラクションが Reject, つまり 2 者のインタラクションが終了してしまうことがある[17]. 報酬を得ることを目標とするインタラクションにおいて, 報酬を得られない or 減らされる場合, インタラクションは終了する可能性があると考えられる. そこで B が報酬を得ていない際は, インタラクションが終了する状態に遷移することを考慮する必要があると考えられる. 各状態とその遷移条件を表 5 (pp.10 に記載) に示す.

### 3. シミュレーション

#### 3.1. シミュレーション目的

報酬量の変化に基づいて、ツンデレインタラクション時の2者の行動及び、コミュニケーション欲求の変化や報酬量の変化をシミュレーションにより調査する。今回の実験では、2者間のみでの実験とし、第三者の介入を考えないものとする。またAの行動に対して報酬を得た・減らされたBは、変化した報酬量にちなんだ振る舞いを行うようにする。それによりAはBの報酬量を予測するものとする。なお今回はシミュレーションで行うため、その予想と実際の報酬量が完全に一致するものとする。

#### 3.2. シミュレーション環境

実験は計算機上に作成するプログラムによってシミュレーションを行う。図3中の三角形は仮想エージェントであり、点線は移動履歴を示している。

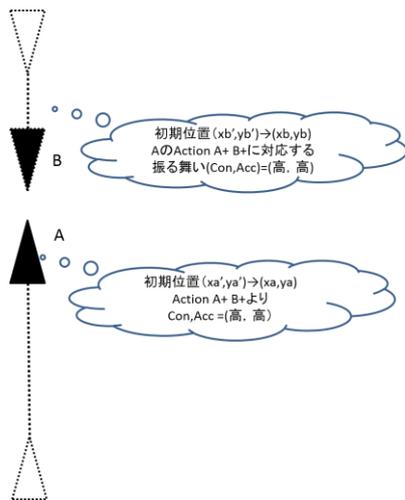


図3:プログラムのイメージ図

#### 3.3. アルゴリズムと手順

ツンデレインタラクションにおいて2者の内部状態の決定と行動、及び報酬による次状態の変化について図4に示す。まずActionを行うためのモードの決定方法を示す。

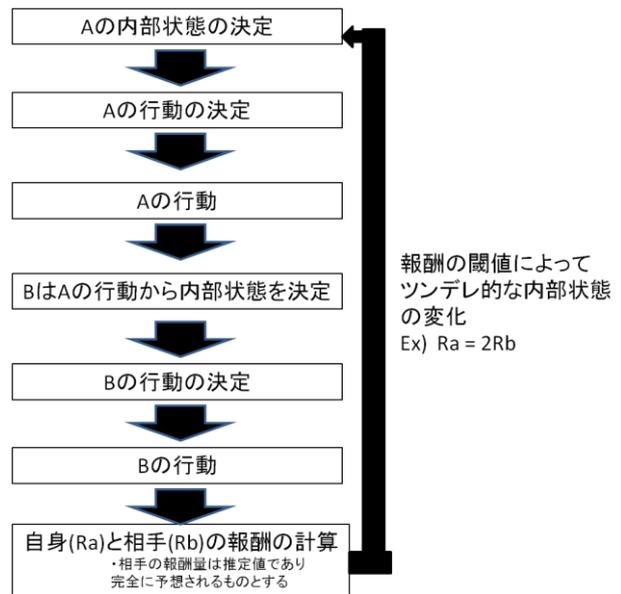


図4:手順のフローチャート

次状態を決定する  $N\_MODE$  は式3によって示される。ツンからデレ、デレからツンなどへの状態遷移は、表5に記した条件と、2者の相対的な報酬量とツンデレを受ける側(B)がどれほどインタラクションを終了したがるか ( $FTA_B$ )に応じて行われる[17]。

$$N\_MODE = V_A * V_B * FTA_B \dots (式3)$$

$V_A * V_B \dots$  AとBの相対的な報酬量の差

次に各状態におけるActionは、前状態( $P\_MODE$ )と現状態( $MODE$ )によって決定される。なお1つの状態に複数のActionが存在する場合、それぞれのActionの発生確率を定め、それに基づき決定される(発生確率を変数Pとおく)。例えばある状態においてAction1とAction2が存在し、それぞれ70%、30%で発生するというふうにする(式4)。なお初期状態のみ、前状態がないため、現状態とPのみで決定される。

$$Action = P_{MODE} * MODE * P \dots (式4)$$

Actionによるそれぞれの報酬量の変化は、Action以前に元々持っていた  $V_{Pre_x}$  とAction時に獲得する  $V_{Act_x}$  の総和によって決定される(式4)。

$$V_x = V_{Pre_x} + V_{Act_x} \dots (式4)$$

また、表2よりActionによってコミュニケーション欲求の値は決定され、Bのコミュニケーション欲求の値はAのコミュニケーション欲求に対応するように決定される。例えばAction A+ B+であればAは互恵的な行動を行い、

BはAの互恵的行動に対応する互恵的行動を行う。各コミュニケーション欲求の対応に関する表を表6に示す[18].

表 6:A と B のコミュニケーション欲求対応表

A	B
互恵(Con,Acc) = (高, 高)	互恵(Con,Acc) = (高, 高)
服従(Con,Acc) = (低, 高)	支配(Con,Acc) = (高, 低)
支配(Con,Acc) = (高, 低)	服従(Con,Acc) = (低, 高)
忌避(Con,Acc) = (低, 低)	忌避(Con,Acc) = (低, 低)

### 3.4. コミュニケーション欲求と行動

- 互恵的であれば、相手の正面に立つように近づく
- 支配的であれば、相手の背後に回る
- 服従であれば、背を向け相手の直線状に移動
- 忌避であれば、相手に背を向け相手から遠ざかる

### 3.5. 状態遷移

ツンデレコミュニケーションでは 2 者の報酬量とコミュニケーション欲求が変化していく連続的コミュニケーションであり、式 2-5、表 1-5 より図 5 (pp.10 に記載) のようになると考えられ、これに基づいてシミュレーションを行う。図 5 中の状態における Action について、例えば S1 の(a1 > b1 | S1)では、前状態が S1 である場合、b1 より a1 の Act が高確率で選ばれることを示す。また状態遷移の 10%や 15%は現状態から、それぞれの確率で遷移する%を示し、90-10\*(s1→s2)%は、初期は 90%で遷移するが、S1→S2 の遷移を繰り返した数\*10 ずつ遷移する確率が下がっていくことを示す。

### 3.6. 観察項目

本実験ではツンデレインタラクションにおける報酬量の変化とその時の内部状態、行動パターンを明らかにするため以下のものを観察項目とする。

- 2者のインタラクション終了時までの報酬量の変化
- 2者のインタラクション終了時までのコミュニケーション欲求の値の変化
- 2者のインタラクション終了時までの2者間の距離

## 4. 結果と考察

### 4.1. 結果

ツンデレインタラクションと(互恵-互恵)、(服従-支配)、(支配-服従)関係のインタラクションにおける 2

者の Con と Acc, 報酬, 2者間の距離に関して結果を記載する。

#### 4.1.1. 内部状態に基づいた行動

A と B のそれぞれの内部状態に基づいた行動の一例を挙げる。各図の赤は A, 青は B を示している

##### 4.1.1.1. 互恵-互恵関係

互いに接近しあい、ある程度近づいたところで停止する。

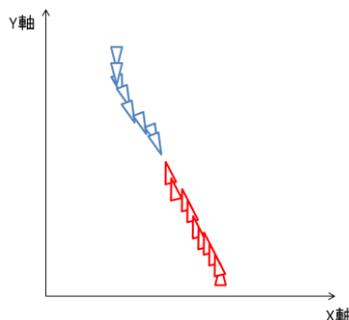


図 6-1 : (A, B) = (互恵, 互恵) の行動パターン

##### 4.1.1.2. 支配-服従関係

B は A に対して背を向け、A は B の背後に移動し停止する。

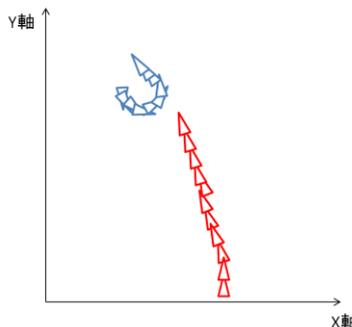


図 6-2 : (A, B) = (支配, 服従) の行動パターン

##### 4.1.1.3. 服従-支配関係

B は A の背後に移動し、A は B に対して背を向け停止する。

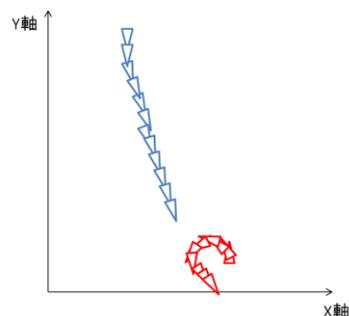


図 6-3 : (A, B) = (服従, 支配) の行動パターン

#### 4.1.1.4. 忌避関係

B が A とのインタラクションを拒み、A に背を向け A から遠ざかる。

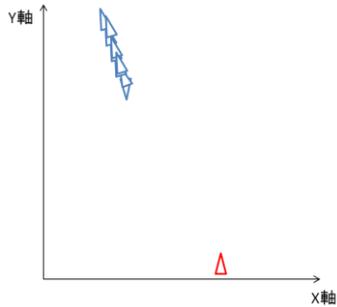


図 6-4 : B が忌避状態の行動パターン

#### 4.1.2. 内部状態と報酬と距離

各内部状態と報酬と距離に関する一例を挙げる。図中赤線は A, 青先は B の変化を示しており, 左上は A と B の Control の変化を時系列的に表している。右上は A と B の Acceptance の変化を時系列的に表している。左下は A と B の報酬の変化を表している。右下は A と B 間の距離を表している。

##### 4.1.2.1. (A,B) = (互恵,互恵) 関係

2 者は図 6-1 のように行動し, A と B の (Con, Acc) は互いに (高, 高) のままインタラクションが続き, 表 2 にあるように互いに報酬が増加していく。しかし互いにある程度近づいたところで停止し, 報酬の変化が停止する。

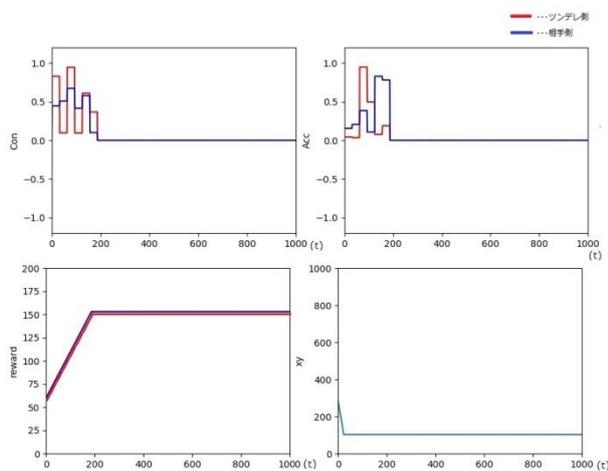


図 7-1 : 互恵関係の内部状態と報酬と 2 者間の距離

##### 4.1.2.2. (A, B) = (支配, 服従) 関係

2 者は図 6-2 のように行動し, A の (Con, Acc) = (高, 低) であり, B の (Con, Acc) = (低, 高) のままインタラクションが続き, 表 2 にあるように A の報酬は増えていくが, B の報酬は減少していく。その後報酬が減少していった B

がインタラクションを終了させようと, 図 6-4 のように行動し, インタラクションは終了し互いの報酬は減少していく。なお B が A から離れていくため, A と B 間の距離は開いていく。

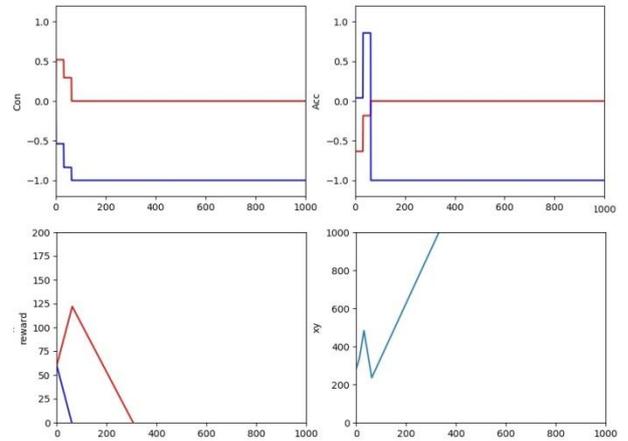


図 7-2 : 支配-服従関係の内部状態と報酬と 2 者間の距離

##### 4.1.2.3. (A, B) = (服従, 支配) 関係

2 者は図 6-3 のように行動し, A の (Con, Acc) = (低, 高) であり, B の (Con, Acc) = (高, 低) のままインタラクションが続き, 表 2 にあるように A の報酬は減少していくが, B の報酬は増えていく。

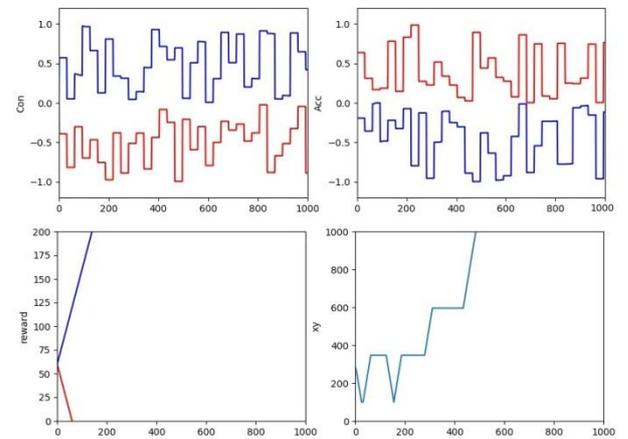


図 7-3 : 服従-支配関係の内部状態と報酬と 2 者間の距離

#### 4.1.2.4. ツンデレインタラクション

##### 4.1.2.4.1. 成功パターン

ツン状態では図 6-2 のように行動し, A の (Con, Acc) = (高, 低) であり, B の (Con, Acc) = (低, 高) のままインタラクションが続き, 表 2 にあるように A の報酬は増えていくが, B の報酬は減少していく。デレ状態では図 6-3 のように行動し, A の (Con, Acc) = (低, 高) であり, B の (Con, Acc) = (高, 低) のままインタラクションが続き, 表 2 にあるように A の報酬は減少していくが, B の報

酬は増えていく。また報酬に関して、x 軸（時間軸）が 250-300 程度と 600-700 程度の間ではゲインロス効果により、本来 B にとって報酬が減るはずのツン状態であるのに報酬が増えている。

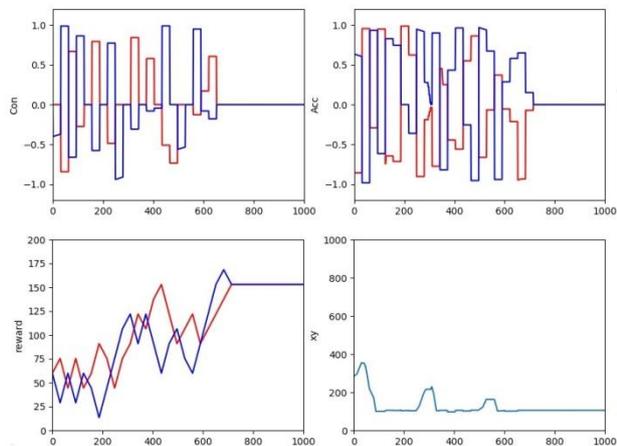


図 7-4：ツンデレインタラクション（成功パターン）

#### 4.1.2.4.2. 失敗パターン

ツン状態とデレ状態における行動と Control と Acceptance, 報酬の増減の仕方は成功パターンに似ている。x 軸が 800 移行から報酬が減少していった B がインタラクションを終了させようと、図 6-4 のように行動し、インタラクションは終了し互いの報酬は減少していく。なお B が A から離れていくため、A と B 間の距離は開いていく。

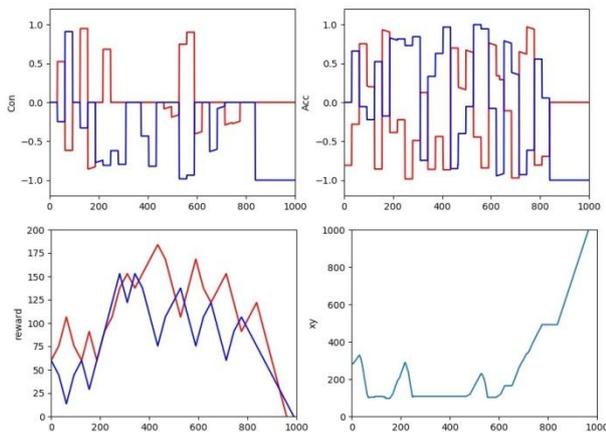


図 7-5：ツンデレインタラクション（失敗パターン）

## 4.2. 考察

(A,B) = (互惠,互惠) 関係では互いに報酬は増えるが、A,B 間の距離は縮まり互いに接近が不可能になる。このとき位置関係のみのインタラクションにおいて、移動が不可能になるため互いの報酬の増加が行われなくなる。これは

初対面の人同士が仲良くなろうとし接近するが、ある程度仲良くなったあとそれ以上仲が進展しない状況に似ていると考えられる。

(A, B) = (支配, 服従) 関係では A, B 間の距離は近づく・離れるを行い位置関係から見たインタラクションの継続は可能である。A の報酬は増え続けるが B の報酬は減り続けるため、B がインタラクションを終了させる可能性が高い。これは横暴な性格の人(A)が B に対してインタラクションを行うが、B が A の横暴な性格に対して嫌気が刺し、B が A から離れていく状況に似ていると考えられる。

(A, B) = (服従, 支配) 関係では A, B 間の距離は近づく・離れるを行い位置関係から見たインタラクションの継続は可能である。B の報酬は増え続け、B がインタラクションを終了させる可能性が低いが、A の報酬は減り続けるため A 自身がインタラクションをするのをやめてしまう可能性が考えられる。これは A が B に対して下手にインタラクションを行うが、B が他者から報酬を受け取ったとき、「恩を返さなければいけない」という返報義務感を生じさせずに A が B に対してインタラクションをするのを諦めてしまうケースに似ていると考えられる。

ツンデレインタラクションにおいて A は基本支配行動を行うが稀に服従行動を行う。その際ゲインロス効果により、A が B に与える報酬が利得の高い報酬として受け止められる。さらに支配と服従を何度も続けることによって、B にとって A にかかわられること辞退が報酬となり、支配行動をされても報酬とを感じるようになったため、図 7-4 のように振動しながら右肩上がりに互いの報酬が増えていったと考えられる。また A, B 間の距離は近づく・離れるを行い、位置関係から見たインタラクションの継続は可能であり、互惠-互惠関係より長期的に互いの報酬を高められるインタラクションの可能性が示唆された。これは人が猫に対して触れようとし近付くと遠ざかるが、遠ざかろうとすると猫側から近づいてきて、最初から近づいてきて触れられる猫よりも、いざ触れられたときの喜びが大きく感じるというインタラクションに似ている。しかし図 7-5 のように A が支配行動をした際に、B の報酬は減り続けるため、B がインタラクションを終了させてしまう可能性が考えられる。これは先程の猫の例でいうところ、猫が人に触れられようとする遠ざかり、人が遠ざかろうとすると近づいてくるを繰り返す、いつまで経っても人が猫に触れないため飽きて去ってしまうというインタラクションに似ている。今後モデルの拡張として、B がインタラクションを終えようとする意図推定を含めたモデル化が必要であると考えられる。

## 5. まとめと今後

我々の日々の生活の中における交渉事、DV など、客観的に見れば一方が一方を苦しめているようだが、本人達には

とっては互恵になるインタラクションがある。本研究では基本的に相手に損をさせるが稀に報酬を与える場合に利得が高く評価されるインタラクションをツンデレインタラクションと定義した。ツンデレインタラクションにおいて、2者の報酬と内部状態の関係性と変化を明らかにし、モデル化することを目的とした。インタラクションに関する知見を参考にして、ツンデレの状態遷移を導出し、実験では、この状態遷移とコミュニケーション欲求を組み合わせた行動パターンに基づいて、シミュレーションを行った。結果として、ツンデレインタラクションは互恵的なインタラクションに比べて、長期的に互いの報酬を高め合うことができることが示唆された。しかし本研究では、他者の報酬量は予測可能であるが、互いの意図推定までは行われなく、ツン状態においてBがインタラクションを終えようとする意図推定を含めたモデルを考える必要がある。今後、一人一人、あるいは一人一物・ロボットでも今回導出されたツンデレインタラクションモデルの行動を行わせたときに、同様の報酬の変化がインタラクションを通じて得られるのかを調査していく。

## 参考文献

- [1]藤澤等, ソシオン理論入門-心と社会の基礎科学-, 北大路書房, 2006
- [2]Walfram Schultz, Neuronal Reward and Decision Signals: From Theories to Data, *Physiological Reviews* 95(3), pp.853-951, 2015.
- [3] Permack, D.& Woodru\_, G: “ Does the Chimpanzee Have a Theory of Mind? ”, *The Behavioral and Brain Sciences* 1, pp.515-523, 1978.
- [4]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応
- [5] Waker, L.E, *The buttered woman*. New York:Harper & Row, 1979
- [6] Aronson, Linder, *The gain-loss principle*,1965
- [7] 坂本孝丈, 竹内勇剛, 須藤明人, 対象との遭遇時における人の内部状態のモデルと接近・回避行動のシミュレーション, HAI シンポジウム 2017, G-22
- [8]Alvin W. Gouldner, *The Norm of Reciprocity: A Preliminary Statement*, *American Sociological Review*, vol.25, no.2, pp.161-178, 1960.
- [9] 野口康彦, 大学生カップル間におけるデートDVと共依存に関する一検討 山梨英, 和大学紀要, 8, 105-113.
- [10] 高野雅典心, 加藤正浩, 有田隆也, 心の理論における再帰のレベルの進化に関する構成論的手法に基づく検討, *Cognitice Studies*, 12(3), 221-223, 2005.
- [11] McCabe, K., Houser, D., Ryan, L., Smith, V.,& Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *PNAS*, 98 (20), 11832—11835.
- [12]高橋泰岳, 河又輝泰, 浅田稔, 自己の価値に基づく他者行動理解, 知能と情報(日本知能情報ファジィ学会誌), Vol.21, No.3, pp.381-391, 2009.
- [13]大坊都夫, 対人コミュニケーションの社会性, 対人社会心理学研究, 1, pp. 1-16, 2001.
- [14]大坊都夫, 親密な関係を移す対人コミュニケーション, 対人社会心理学研究, 4, pp.1-10, 2004.
- [15] Clark, M. S. & Mills, J. 1979 Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology*, 37, 12-24.
- [16]湯浅将英, 安村禎明, 新田克己, 主観的要素を考慮した交渉の状態遷移モデル, システム制御情報学会論文誌, Vol.14, No.9, pp.439-446, 2001.
- [17]宮本友樹, 片上大輔, 重光由加, 宇佐美まゆみ, 田中貴紘, 金森等, ポライトネス・ストラテジーに基づく会話エージェントとの言語的な振る舞いの違いが人との関係性構築にもたらす効果～初対面における冗談の心理効果～, 知能と情報, 日本知能情報ファジィ学会誌, Vol.30, No.5, pp.753-765,2018.
- [18] *Analysis of Relational Communication in Dyads: New Measurement Procedures*, Roger L Edna, Farace Richard, Annual Conference of the International Communication Association, Montreal, April, pp25-29, 1973.

表 5 : 各状態と遷移条件

状態	遷移先1	条件1	遷移先2	条件2	遷移先3	条件3	
S1	ツン1	ツン1(S1)	2A<B	デレ(S2)	2A≧B	Se(終了)	一定確率
S2	デレ	デレ(S2)	A>B	ツン1(S1)	A≧B(S1→S2回数が少ないほど遷移する)	ツン2(S3)	A≧B(S1→S2回が多いほど遷移する)
S3	ツン2	ツン2(S3)	S3→S3回数が少ないほど遷移する	ツン3(S4)	S3→S3回が多いほど遷移する		
S4	ツン3	S2(デレ)	Seに遷移しなければ	Se(終了)	一定確率		
Se	終了						

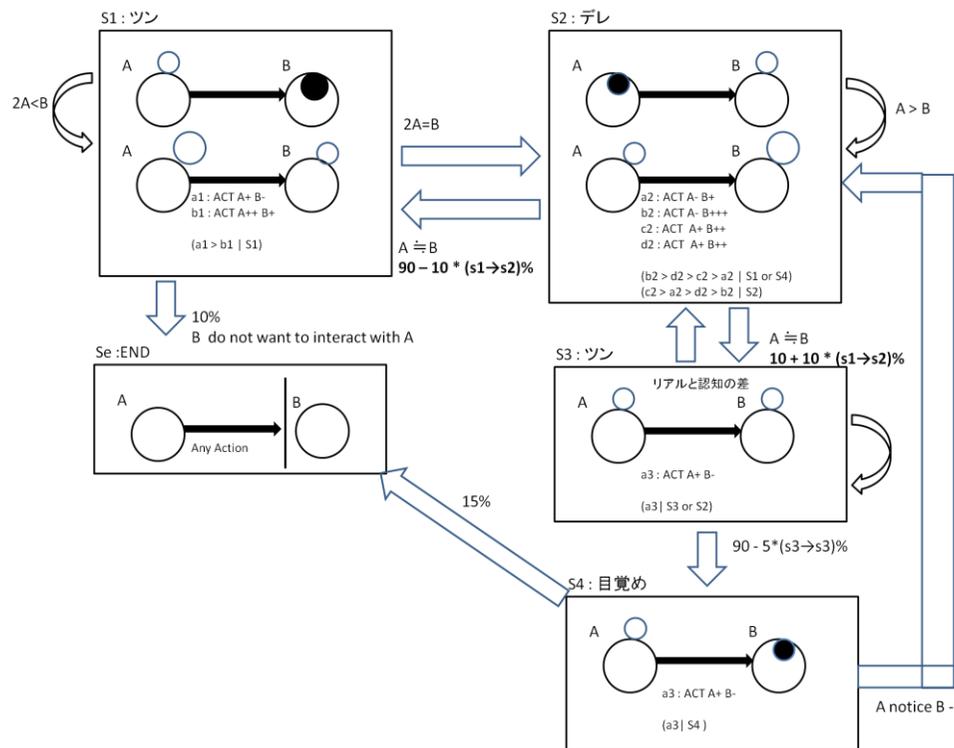


図 5 : ツンデレインタクション