

# CA モデルに基づく狭路での人同士のすれ違いのシミュレーション

## Simulation of passing each other on narrow roads based on CA model

江原 脩<sup>1\*</sup> 竹内 勇剛<sup>1</sup>

Osamu EBARA<sup>1</sup> and Yugo TAKEUCHI<sup>1</sup>

<sup>1</sup> 静岡大学大学院総合科学技術研究科 〒432-8011 静岡県浜松市中区城北3丁目5-1

<sup>1</sup>Department of Informatics, Shizuoka University 3-5-1

Johoku, Naka-ku, Hamamatsu, Shizuoka, 432-8011 Japan

**Abstract:** 近年、「CarriRo Deli」といった自律走行を行うロボットの存在が注目を集めている。しかし、自律走行型ロボットの普及の見通しはまだ立っていない。その理由としては自律走行型ロボットと人とのインタラクションデザインが未だ実用段階にないためである。自律走行型ロボットの実用場面を考えた場合、狭路におけるすれ違いの問題は避けては通れない。この問題を解決するためには、狭路でのすれ違い場面における人の認知過程を解明し、ロボットに理解させる必要がある。そこで本研究では CA モデルと呼ばれる人の内部状態モデルに基づいた行動を生成する 2 体のエージェントを用いてすれ違いシミュレーションによる強化学習によって最適なすれ違い行動を学習させ、実際の人のすれ違い行動と比較することで狭路でのすれ違い場面における人の認知過程を明らかにする。これによって、人の対人インタラクションにおける特徴を CA モデルに基づいて明らかにし、自律走行型ロボットの走行デザインへの指針を与える。

### 1 はじめに

近年、「CarriRo Deli」といった自律走行を行うロボットの存在が注目を集めている。今後ロボットの研究が進み、人と同じ生活空間において多くの自律走行ロボットが活躍することが期待される。自律走行を行うロボットの実用場面を想定した場合、狭路におけるすれ違いの問題は避けては通れない。具体的には人通りの多い歩道などでは、ロボットが人をうまく回避できず同じ方向に避けてしまうお見合いや人の影に隠れて突然出てきた対向者を避けきれずに衝突する場合が発生することが考えられる。

この問題に関する研究のアプローチとしては 2 種類存在している。一つはロボット自体にセンサーを取り付け、周囲の環境を認識し、すれ違いを行う方法、もう一つは環境に取り付けたセンサー情報を利用し、すれ違いを行う方法である。前者に関してはパーソナルスペースの観点から人のすれ違い時の動きをモデル化し、ロボットに取り付けたセンサーから行動を予測し、すれ違いを実現した研究 [1][2] や、カメラから得られる対向者が回避する際の顔や胴体、足の向きといった動きのデータ

から回避方向を予測し、すれ違いを実現した研究 [3] など人の動作に着目した研究が多く見られる [4][5]。一方で、環境に取り付けたセンサー情報を用いたアプローチでは環境内に配置されている複数の CCD カメラを用いて、空間を歩行する複数人をトラッキングすることで行動を予測し、すれ違いを実現する研究 [6][7] や各時間帯における人の頻度をグリッドマップとして蓄積し、ロボットの経路計画に利用することによってすれ違いを実現する研究 [8] など人個人の動作ではなく全体の人の流れに着目している研究が多くみられる。いずれのアプローチにおいてもすれ違いという動作に関しては一応の実現はなされている。しかし、どちらのアプローチ方法でも回避時のお見合いといった問題の根本的な解決法にはなっていない。その理由としては行動生成時に人の欲求や感情といった心理的情報について考慮されていないことがあげられる。

人には欲求というものが存在し、無自覚的に自分の欲求に基づきその効用を最大化、もしくは他者の欲求に基づく効用を最大化する振る舞いをしてしまう [9]。従来の工学的なアプローチでは、センサー等による物理的な情報のみを取得し、行動を生成しており、人の欲求や感情といった心理的情報が考慮されていない。そのため、一見して最適に見えるすれ違い動作であってもそれは

\*静岡大学大学院総合科学技術研究科情報学専攻  
〒432-8011 静岡県浜松市中区城北3丁目5-1

物理的な計算によって求められた最適解であり、人同士で行っているような心理的な計算が含まれていないためお見合いのような衝突が生まれてしまっていると考えられる。この問題を解決するためには、狭路のようなすれ違いが起きてしまう場面における人の認知の過程を解明し、ロボットに理解させる必要がある。

そこで、本研究では坂本ら [10] が提案している CA モデルと呼ばれる接近・回避行動モデルを参照し、CA モデルに基づいた行動を生成する 2 体のエージェントによるすれ違いシミュレーションを行う。その際、片方のエージェントの内部状態を固定した上で ML-Agents による強化学習を行うことによって相手の内部状態に応じたすれ違い動作を学習させ、そこで得られたエージェントの軌道と実際に人がすれ違う際に描く軌跡とを比較することで狭路でのすれ違い場面における人の認知過程を CA モデルに基づき明らかにする。これによって、人の認知に負荷を与えずに効率的にすれ違いを実現するための行動を計算論的に明らかにし、自律走行型ロボットの走行デザインへの指針を与える。

## 2 背景

### 2.1 CA モデル

人には「心の理論」と呼ばれる他者の心的状態を類推する機能が備わっており [11]、他者と何かしらのインタラクションを行う際には、この「心の理論」を用いることによって相手の行動意図や感情といった内部構造を推測し、自身の行動を決定しているとされている [12][13]。

すれ違いという行動に関しても同様の駆け引きが自身の相手との間で行われていると考えられる。例えば、接近してくる相手に対して回避運動を行う場合、相手の目線やつま先の方から相手がどの方向へ向かおうとしているかを推測することで回避の方向を決定、相手の歩行速度から相手の心理状態を推測することでその行動をどのタイミングで行うかを決定しているのは日常的な経験として想像に難くない。このように、人の行動と内部状態というのは密接なかかわりを持っている。

この内部状態と行動との関係を示したモデルに CA モデルと呼ばれる接近・回避行動モデルがある。CA モデルではエージェントの内部状態と相手との空間的な配置から行動が生成されると同時に接近や回避行動から相手の内部状態を推測することが可能なモデルである。

すれ違いの行動についても CA モデルに基づき行動を生成したならば、相手の反応から内部状態を推測し、相手の認知的負担を最小限に抑えたすれ違いを可能にする行動モデル生成できるのではにかと考える。

従って本研究では CA モデルに基づきすれ違い行動を生成するエージェントを用いてすれ違いシミュレーションによる強化学習を行い、エージェントに最適なすれ違いを実現する内部状態を学習させることによって人の認知的負担を最小限に抑えるための効率的なすれ違いを実現する行動を CA モデルに基づいて計算論的に明らかにする。そのために以下より、CA モデルで取り扱っている内部状態と行動についての詳細を説明する。

#### 2.1.1 CA モデルにおける内部状態

CA モデルにおいては、あるエージェントから見た相手エージェントとの関係は自身から相手に関与しようとする選好 (*Control*) と相手から自身への関与に対する選好 (*Acceptance*) という 2 軸によって記述できるとされ、*Control* の値を  $c$ 、*Acceptance* の値を  $a$  としたときにエージェントの内部状態は  $z = (c, a)$  の形で記述される。内部状態  $z$  は大まかに *Control* と *Acceptance* の正と負の組み合わせた 4 つの状態と選好を持たない状態の 5 つの状態に分類される (図 1)。この 5 つの状態については以下のように定義されている

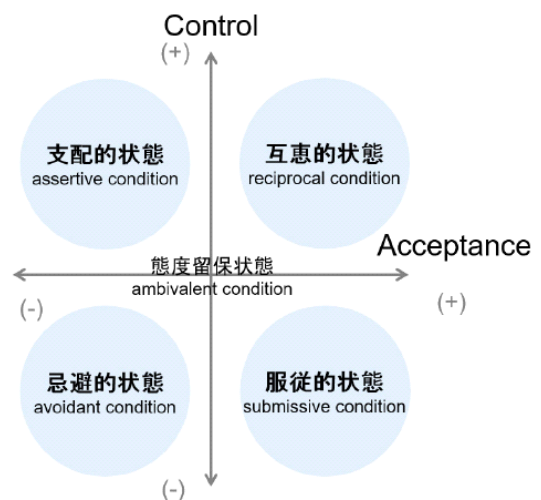


図 1: CA モデルにおける内部状態 [10]

- $c > 0, a > 0$ : 自身と相手が共に関与可能な状況を選好する状態 (互惠状態)
- $c > 0, a < 0$ : 自身から相手への関与のみが可能な状況を選好する状態 (支配状態)
- $c < 0, a > 0$ : 相手から自身への関与のみが可能な状況を選好する状態 (服従状態)
- $c < 0, a < 0$ : 自身と相手が共に関与を拒む状況を選好する状態 (忌避状態)

- $c = 0, a = 0$  : 自身と相手が共に関与可能な状況を選好しない状態 (保留状態)

### 2.1.2 関与度と空間的配置

内部状態と実空間上の配置との対応付けとして, CAモデルでは人による対象への関与の強さ  $\alpha$  は身体の相対的な位置関係のうち, エージェントと対象間の距離  $r$  と相対角度  $\theta$  によって定義され, 以下の式で表される.

$$\alpha = f(r, \theta) \quad (1)$$

ただし,  $f(r, \theta)$  は単調減少関数であり,  $\alpha > 0$  と定義されており, エージェントが行動を生成するために用いる相対的な位置関係は距離  $r$  と2つの相対角度  $\theta_{ij}$ ,  $\theta_{ji}$  は次のようように定義される (図2).

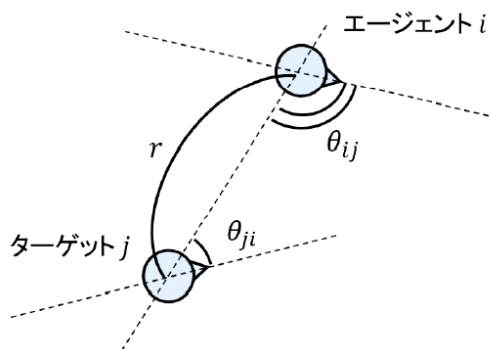


図2: 関与度と空間配置の関係 [10]

### 2.1.3 行動の生成

先ほどの内部状態の定義と空間的配置との関係から CAモデルでは図3で示されるように生成する行動は平面方向への方向行動と身体の回転のみを取り扱っており, それぞれの移動ベクトルは次のように定義されている.

$$\Delta_i r = v_{dis}(c + a, r) \quad (2)$$

$$\Delta_i \theta_{ji} = v_{rev}(a, \theta_{ji}) \quad (3)$$

$$\Delta_i \theta_{ij} = v_{turn}(c, \theta_{ji}) \quad (4)$$

$v_{dis}, v_{rev}, v_{turn}$  はそれぞれ相手との距離方向による成分, 相手の顔の向きに対する回転成分, 自身の身体方向に対する回転成分である.  $v_{dis}$  は自身と相手の関与の強さによって変化し, 相手との距離  $r$  を縮めることで大きく, 離れることで小さくなる, また,  $v_{rev}, v_{turn}$  は自身から相手, 相手から自身への関与の強さによって変化し,  $v_{rev}$

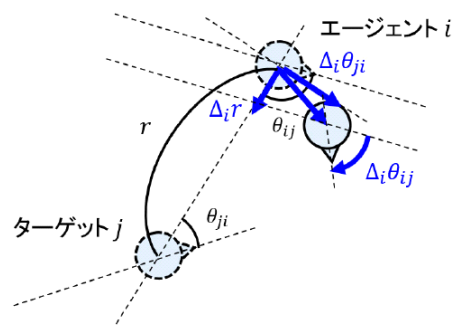


図3: エージェントの接近・回避行動を構成する成分 [10]

は自身と相手との相対角度  $\theta$  が大きくなるほど,  $v_{turn}$  は自身と相手との相対角度  $\theta$  小さくなるほど値が大きくなる.

## 2.2 ML-Agents による強化学習

本研究では狭路においてエージェントに最適なくすれ違い動作を獲得させるにあたって ML-Agents による強化学習によってすれ違い動作の学習を行う.

一般的な自律型ロボットやエージェントの行動決定にはあらかじめロボットが動作する環境やそれに伴う影響について厳密な数学モデルによって予測し, 行動を計画しロボット側に組み込む必要がある. しかし, 強化学習を用いた場合, これらの複雑な数学モデルを必要とせず, 自身の試行錯誤によって獲得される学習データから目的を達成するための行動計画を自ら学習することができる. 現に自律型ロボットの行動獲得において強化学習や深層学習を用いてロボットに目的達成のための行動を獲得させる研究というものがすでに数多く為されている [14, 15]. このように, 自律的行動を獲得させる手法として強化学習を用いることは有効な手段である. しかし, 一般に強化学習には膨大な学習データを必要とするため, 実機を用いたデータ収集には莫大なコストがかかってしまうという問題がある.

そこで, 本研究では Unity 内に搭載されている ML-Agents と呼ばれる強化学習機能を用いてシミュレーションを行う. ML-Agents とは Unity 内に用意されている強化学習機の一つである. Unity には 3D ゲーム開発を行うための統合開発環境が整っており, Unity を用いることによって様々な現実の環境状態を疑似的に再現することが可能である.

本研究においては仮想空間上において CAモデルに基づいた行動を生成する Agents を2体作成し, 仮想空間上の狭路において互いに向き合った状態で前進させ, すれ違いを行わせることによって最適なすれ違いを行

う CA モデルを学習させる。学習の際、片方の Agent の CA モデルを固定し、もう片方の Agent に ML-Agents によって最適な CA モデルを学習させることによってすれ違いを行う相手が持つそれぞれの CA モデルに対応した CA モデルの取得を行う。取得した CA モデルによって生成されるすれ違い行動の軌道と現実の人のすれ違い行動の軌道とを比較・分析を行うことによってすれ違いにおけるインタラクションの特徴を CA モデルに基づいて明らかにする。

### 3 CA モデルに基づくすれ違いシミュレーション

本研究では狭路でのすれ違い場面における人の認知過程を明らかにすることを目的に、坂本らの CA モデルの定義に基づきすれ違い動作を生成するエージェントを作成、強化学習によるすれ違いシミュレーションを行わせ、エージェントに最適なすれ違いを実現する内部状態を学習させることによって人の認知的負荷を最小限に抑えるための効率的なすれ違いを実現する行動を CA モデルに基づいて計算論的に明らかにする。以下では、すれ違いシミュレーションにおいて取り扱うエージェントの行動の生成過程と内部状態の学習方法について説明する。

#### 3.1 CA モデルに基づくすれ違い行動の生成

シミュレーションにおいてエージェントが生成するすれ違い動作については CA モデルの距離成分  $v_{dis}$ 、回転成分  $v_{rev}$  に加えて、すれ違った先の目標地点へ向かう移動成分  $v_{tar}$  を含めた 3 つのパラメータ成分によって定める、それぞれの移動ベクトルと空間配置は次のように表す(図 4~7)。

- すれ違った先にある目標地点への移動成分  $V_{tar}$
- 相手との距離に基づく移動成分  $V_{dis}$
- 相手の向きに基づく回転成分  $V_{rev}$

移動ベクトル  $V$  はこれら 3 つのベクトル成分の合成によって方向を決定する。その際、目標地点へ向かう意思の強さと相手への関心の度合いによって行動が変化すると考えられるため、各成分に重みづけ ( $w_1, w_2, w_3$ ) を行う。加えて、一般的なすれ違いの場面においては目標地点へ向かう意思のほうが相手への関心よりも強くなると考えられるため、3 つの重みの大小関係は  $w_1 > w_2, w_3$  のようになると考えられる。以上のことを踏まえ、生成される移動ベクトル  $V$  を下記に示す。

$$V = w_1 * v_{tar} + w_2 * v_{dis} + w_3 * v_{rev} \quad (5)$$

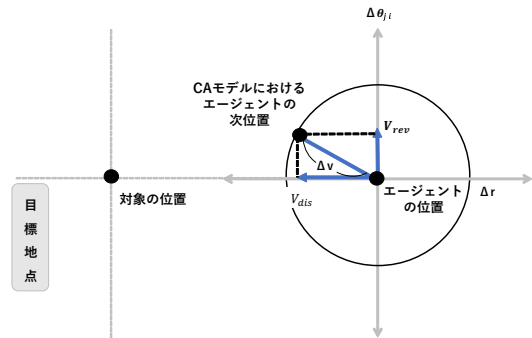


図 4: CA モデルによる構成成分 (モデル)

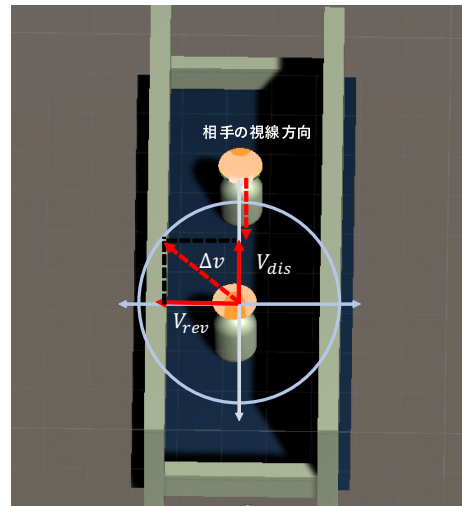


図 5: CA モデルによる構成成分 (シミュレーター上)

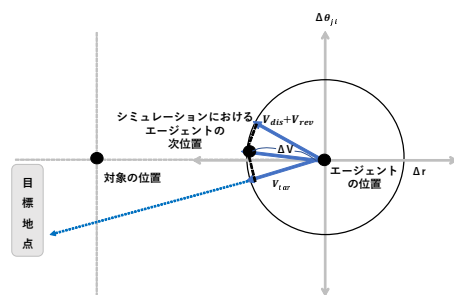


図 6: すれ違い動作を構成する成分 (モデル)



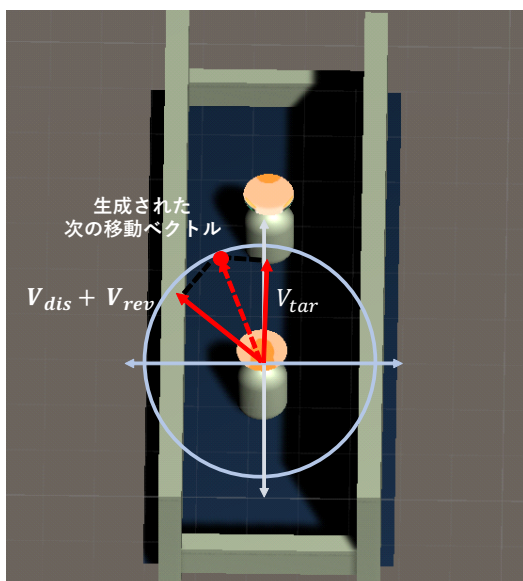


図 7: すれ違い動作を構成する成分 (シミュレーター上)

### 3.2 学習における観測値と報酬

*ML - Agents* では方策勾配法を採用している。方策勾配法では学習者が環境から得た情報をもとに目的達成に向けた行動を行い、それによって得られた報酬を学習データとして得られる報酬を最大化する行動を確率的に計算し、最適な方策の学習を行っている。

本シミュレーションでは、エージェントに相手の内部状態に応じたすれ違い動作を学習させるにあたって、相手の内部状態を推測するために必要な自分から相手に対する相対角度  $\theta_{ij}$ , 相手から自分に対する相対角度  $\theta_{ji}$ , 相手との相対距離  $r$  の 3 つの値に加えて、目的地へたどり着くために目的地までの相対距離  $r_{goal}$  を観測値として与え、報酬には、目的地にたどり着いたならば最大報酬 1.0 を、相手と衝突してしまったならば最低報酬 -1.0 を与えるよう設計を行った。また、相手の内部状態に応じたすれ違い動作を学習させるにあたって、下記に示す効用関数を定義する。

$$U = w_4((1 - \alpha_{ij}) + (1 - \alpha_{ji})) + w_5\left(\frac{1}{\cos_i} + \frac{1}{\cos_j}\right) \quad (6)$$

CA モデルより  $\alpha_{ji}$ ,  $\alpha_{ji}$  は自分から相手、もしくは相手から自分に対する関与の強さを表し、 $\cos_i$ ,  $\cos_j$  はすれ違い行動によって自分または相手が支払ったコストを表す。ここでのコストはすれ違いが完了するまでに費やした時間と移動距離とする。人込みや一対一といった状況を想定した場合、人込みのような一人一人に関心を向けられない状況においては相手への関与の度合いより費やすコストの方を優先した行動を行うと考えられる。逆に、一対一のように相手に対して関心を向けなくてはならない状況においては相手の関与の度合いの方が費やすコストよりも優先した行動を行うと考えられ

る。

従って、それらの状況を設定するために各成分に対して重み付け ( $w_4, w_5$ ) を行い、それぞれの状況において効用関数  $U$  を最大にするような  $V$  を学習させる。

### 3.3 シミュレーション条件

シミュレーションによる強化学習を行う際、片方のエージェントの CA パラメータの値を固定して行い、もう一方のエージェントの内部状態を自由に遷移させた上ですれ違い動作を学習させることで、各 CA モデルの状態の相手に対して最適なすれ違いを実現するためにはどのような内部状態遷移が必要とされるのかシミュレーションを通じて検証する。

## 4 シミュレーション結果

学習によって各 CA モデルの内部状態に対するすれ違い行動による軌跡を以下に示す (図 8~11) なお、軌跡については青が学習した Agent の軌道、橙が内部状態を固定された Agent の軌道である。また、その際の Agent の内部状態遷移のグラフを以下に示す (図 12~15)。

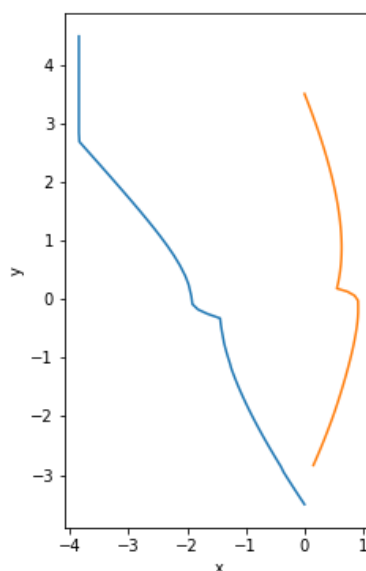


図 8:  $C=1, A=1$  の内部状態を持つ相手に対するすれ違い軌道

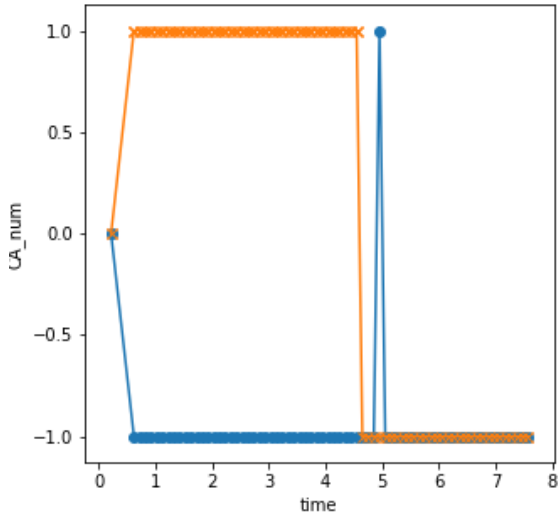


図 9:  $C=1, A=1$  の相手とのすれ違い時の内部状態遷移 (青: Ctrl、橙: Acc)

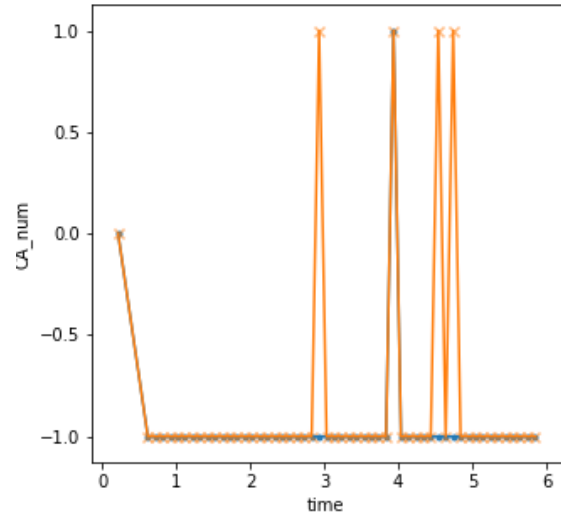


図 11:  $C=1, A=-1$  の相手とのすれ違い時の内部状態遷移 (青: Ctrl、橙: Acc)

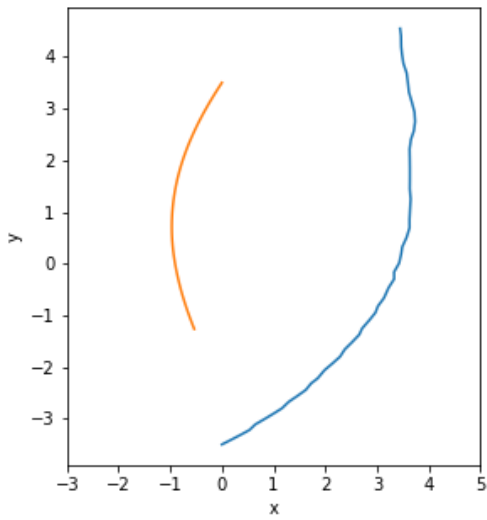


図 10:  $C=1, A=-1$  の内部状態を持つ相手に対するすれ違い軌道

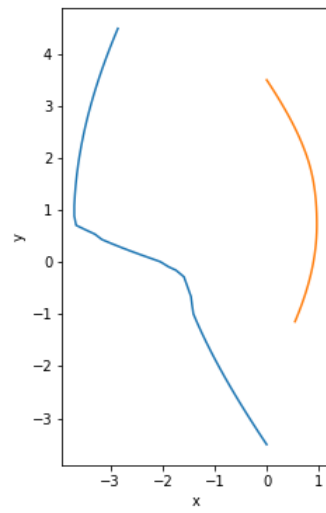


図 12:  $C=-1, A=1$  の内部状態を持つ相手に対するすれ違い軌道

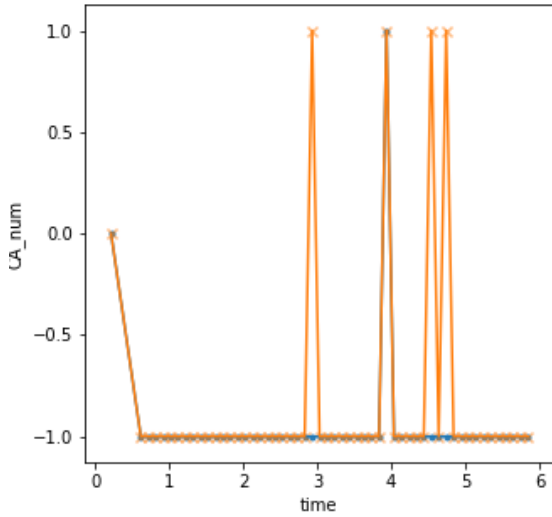


図 13:  $C=1, A=-1$  の相手に対するすれ違い時の内部状態遷移 (青: Ctrl、橙: Acc)

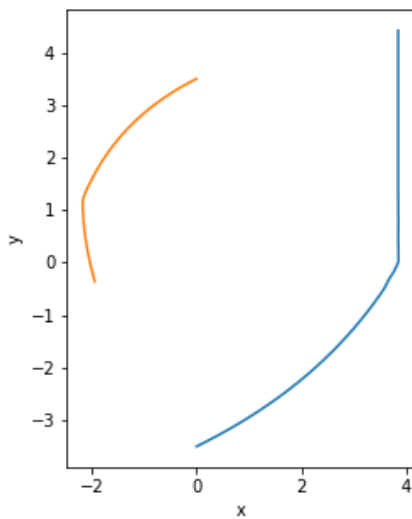


図 14:  $C=-1, A=-1$  の内部状態を持つ相手に対するすれ違い軌道

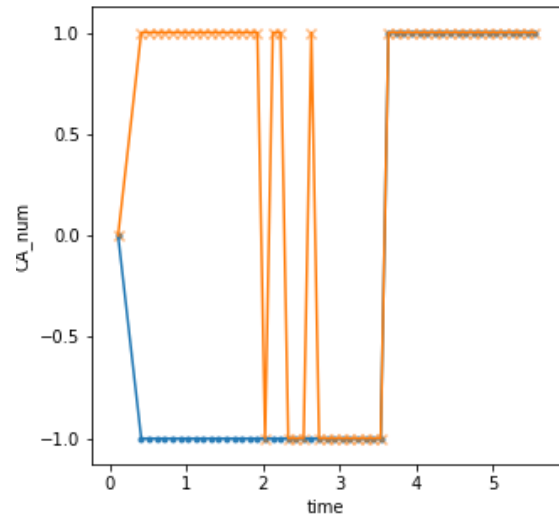


図 15:  $C=-1, A=1$  の相手に対するすれ違い時の内部状態遷移 (青: Ctrl、橙: Acc)

## 5 考察

得られたすれ違い動作による軌跡と内部状態の遷移の様子から、すれ違う際は主に Acc を変動させることによって行動を決定し、Ctrl に関しては相手がいかなる状態においても常に負の状態にしておき、相手の行動から、相手の Acc が正である、つまり、相手の方が自身の行動に対して寛容であると推測されたならば自身の Acc 値も活性化させ、逆に相手の Acc 値が非活性ならば自身も非活性状態に遷移する行動が学習された。これは、初めから相手に行動の主権を握らせるのではなく相手側が行動の主権を放棄したと行動から判断した場合に自身の Acc を下げ、相手と同じ立場になろうとしているのではないかと考えられる。

## 6 まとめと今後の計画

本稿では、近年、自動運転車などで話題となっている自律走行ロボットの現状と抱えている問題について取り上げ、狭路でのすれ違いにおける問題に対して CA モデルと呼ばれる人の内部状態モデルに基づいた行動を生成する 2 体のエージェントを用いたすれ違いシミュレーションによる強化学習を行うことで最適なすれ違い動作の学習を行った。

その結果、すれ違い行動時においては自身の内部状態の Ctrl 値を常に非活性状態にしておき、相手の行動に応じ、Acc の値の変動によってその行動を遷移させる行動が学習された。また、Acc 値の変動については相

手の行動から、相手の Acc 値が活性状態と推測されたならば自身の Acc 値も活性化させ、逆に相手の Acc 値が非活性ならば自身も非活性状態に遷移する行動が学習された。今後の計画としては今回得られたすれ違い軌道と実際の人がすれ違いの際に描く輝度とを照らし合わせ、どれだけ一致しているのか追加実験によって検証を行う予定である。

## 7 謝辞

はじめに、本研究に際して様々なご指導を頂きました竹内勇剛教授に深謝いたします。そして多くのご指摘をくださいました竹内研究室の先輩・同期の皆様方、そして、実験の際に実験協力を快く引き受けてくださった先輩・同期・後輩や、学部学科の友人など、皆様へ心からの感謝の気持ちと御礼を申し上げたく、謝辞にかえさせていただきます。

## 参考文献

- [1] 依田光正, & 塩田泰仁. (1999). 作業の教示とプログラミング 人間とすれ違い行動を行う移動ロボットの研究. 日本ロボット学会誌, 17(2), 202-209.
- [2] 建部謙治, & 中島一. (1990). 静止した障害物に対する単独歩行者の回避行動: 歩行者の回避行動に関する研究 (I). 日本建築学会計画系論文報告集, 418, 51-57.
- [3] 伊藤明久, 今井順一, 金子正秀. 対向者の歩行動作の観測と予測に基づくロボットの自律的なすれ違い. 第7回情報科学技術フォーラム, No. j-049, pp. 511-512, 2008
- [4] 寺田善貴, 田村雄介, 山下淳, & 浅間一. (2012). 移動ロボットの軌道生成のためのすれ違いにおける歩行者の行動パターンのモデル化. 第24回自律分散システムシンポジウム予稿集, 175.
- [5] Thompson, S. , Horiuchi, T. , & Kagami, S. (2009, February). A probabilistic model of human motion and navigation intent for mobile robot path planning. In 2009 4th International Conference on Autonomous Robots and Agents (pp. 663-668). IEEE.
- [6] Morioka, K. , & Hashimoto, H. (2004). Appearance based object identification for distributed vision sensors in intelligent space. In 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 199-204). (2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Vol. 1).
- [7] Lee, J. H. , & Hashimoto, H. (2002). Intelligent space—concept and contents. Advanced Robotics, 16(3), 265-280.
- [8] 本仲君子, 前山祥一, 渡辺桂吾. Hfm を用いた人の移動を妨げない移動ロボットの経路計画. 自動制御連合講演会講演論文集, Vol. 53, pp. 241-241, 2010.
- [9] Parsons, T. , Shils, E. A. , & Smelser, N. J. (Eds. ). (1965). Toward a general theory of action: Theoretical foundations for the social sciences. Transaction publishers.
- [10] 坂本孝丈, 吉岡源太, & 竹内勇剛. (2019). 話しかけ場面における相手の受容度に応じた接近行動のモデルに基づく分析. 知能と情報, 31(5), 842-851.
- [11] Premack, D. , & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. Behavioral and brain sciences, 1(4), 515-526.
- [12] 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
- [13] Goldman, A. I. (1989). Interpretation psychologized. Mind & Language, 4(3), 161-185.
- [14] 山田和明, 大倉和博, & 上田完次. (2003). 強化学習による自律型アームロボットの協調行動獲得. 計測自動制御学会論文集, 39(3), 266-275.
- [15] 宮本賢良, 渡邊紀文, & 武藤佳恭. (2019). すれ違い協調行動のエージェントシミュレーションによるメタ戦略の学習. In 日本知能情報ファジィ学会 ファジィシステムシンポジウム 講演論文集 第35回ファジィシステムシンポジウム (pp. 765-768). 日本知能情報ファジィ学会.