

深層強化学習による車椅子の動作シミュレーションに関する研究

A Study on Simulation of Wheelchair Motion by Deep Reinforcement Learning

○金澤 光^{*1}, 加戸 啓太^{*2}

Hikaru KANAZAWA^{*1} and Keita KADO^{*1}

^{*1} 千葉大学 大学院融合理工学府 博士前期課程

Graduate Student, Graduate School of Science and Engineering, Chiba University

^{*2} 千葉大学 大学院工学研究院 助教 博士(工学)

Assistant Professional, Graduate School of Engineering, Chiba University, Ph.D.

キーワード：深層強化学習、車椅子、バリアフリー、シミュレーション

Keywords: Architecture; structure; environment; planning; information.

1. はじめに

現在、深層強化学習は機械学習や人工知能分野に関する研究など、様々な分野で活用されており、注目を集めている。建築分野においては、自律搬送機械などの制御情報を求めようとする研究が報告されている。藤澤ら¹⁾は深層強化学習を用い、実在建築物における自律搬送機械の運搬シミュレーションを行い、具体的な制御を入力せずとも最適な動作で指定した地点へ到達可能であることを報告している。また、藤岡ら²⁾は、五軸加工機内部における加工間のツールの動作を、深層強化学習を用い求める手法を報告している。これらは従来、人の手により事前に用意していた制御情報を深層強化学習により効率的に得ようとするものである。深層強化学習は、任意の場面における判断を学習することもでき、高添ら³⁾はドローンによる対象に屋内の巡回を目的に、自動導出した経路の巡回中における障害物回避を深層強化学習により実現している。

ゲームの分野では、深層強化学習を用いてゲームのプレイヤーをシミュレーションすることも行われている。ゲーム空間内において、指定した目的、例えば指定した地点にキャラクターを到達させる、を実現するための様々な操作を、深層強化学習を用いた学習において行うことで、ゲーム開発者が意図しない動作においても正常な動作を続けるかを確認する。

本研究では、この深層強化学習をゲームのデバッグに用いることに着目し、建築空間内における人の動作をシミュレーションすることで有益な情報が導けないかと考えた。この報告では、深層強化学習も住宅を模した仮想空間において車椅子を指定した地点まで到達させるタスクを学習させることを行い、その傾向を確認しつつ、その住宅のバリアフリー度合の評価などへの活用できないかということとを考察する。

2. 研究概要

本研究では統合開発環境 Unity^{注1}が提供する強化学習用のフレームワーク ML-Agents^{注2}を用い、住宅空間にて車椅子を指定した地点まで到達させるタスクの学習を行う。これを通し、第一に、このようなタスクを深層強化学習によって行った場合の能否や導出される経路の傾向を観察する。第二に、タスクの難度を変更した場合の能否や安定的にタスクが達成されるまでに要する試行回数から、バリアフリーの程度のような指標が得られないか考察を行う。

2.1. 深層強化学習の概要

深層強化学習とは、強化学習とディープラーニングを合わせた学習手法である。エージェントと呼ばれる主体を介して、環境の状態を観察し、それをポリシーと呼ばれる関数に入力し、出力された行動を実行する。これを繰り返した結果として、タスクを達成すれば報酬を受け取る。試行錯誤を通して将来得られる価値を最大化するようポリシーが最適化される、という仕組みである。以下のように用語を整理する。

- ・ エージェント：行動を行う（あるいは制御を受けて動作をする）主体。本研究においては車椅子が「エージェント」である。
- ・ 環境：エージェントがいる世界。本研究では住宅を模した仮想空間が「環境」である。
- ・ 行動：エージェントが環境内で行う動作。本研究では車椅子の駆動輪（後輪）それぞれに対し前進・後進方向へと回転させようとするトルクを「行動」とした。
- ・ 観察：行動の決定において、環境から得られる入力。エージェントの初期位置およびゴールを含むようにキャプチャした上部からの映像を「観察」とした。
- ・ 報酬：一連の環境（観察）に対する行動の良さを示

す指標。深層学習の教師ありタスクなどでは、入力に対する正解が教師データとして与えられるが、深層強化学習においては、観察（入力）に対する行動の正解は直接与えることは行わず（行えず）、エージェントが環境（観察）に対し、そのような行動をすれば報酬が最大値されるかを学習する。

- ・ ポリシー：観察を入力に行動を出力する関数。
- ・ エピソード：エージェントが環境内に生成されてから、ゴールに到達するなど一連の行動を取り終え終了するまでを指す学習の1回分の単位。
- ・ ステップ数：行動を行った回数。

3. 学習環境の設定

Unity を用い環境となる住宅空間、エージェントである車椅子といった学習環境を用意する。深層強化学習には ML-Agents を用いる。

3.1. 車椅子（エージェント）

深層強化学習のチュートリアルでしばしばエージェントとして用いられる球や自動車は、XY 軸に対する加速度や、ステアリング角などが行動として採用されるが、本研究では車椅子の細かな挙動を反映させることを目的に、図1のように車椅子の機構を模したエージェントを用意した。行動は、駆動輪へのトルクとし、駆動輪の回転によって空間内を走行する。駆動輪の一方を前転、もう一方を後転するよう行動させるとその場で回転する。また、自走式車椅子の前輪は一般に自在車（自在キャスター）である。旋回軸と前輪の回転軸にオフセットがあり、トレイルと呼ばれる旋回軸とタイヤ接地点の距離により曲がる方向に向かって車輪が旋回するようになっている。エージェントの前輪にもトレイルを設け、実際の車椅子の挙動から逸脱しないよう心掛けた。寸法は高さ 870mm x 奥行 900mm x 幅 700mm とした。

エージェントには、車椅子を囲むように透明な直方体（ここでは衝突判定ボックスと呼ぶ）を設け、これに壁やゴールオブジェクトなどとの当たり判定を付与した。このボックスの大きさをいくつかのパターンに変化させることで、間取りにおける開口などの有効寸法と、車椅子におけるバリアフリーの関係について考察していく。

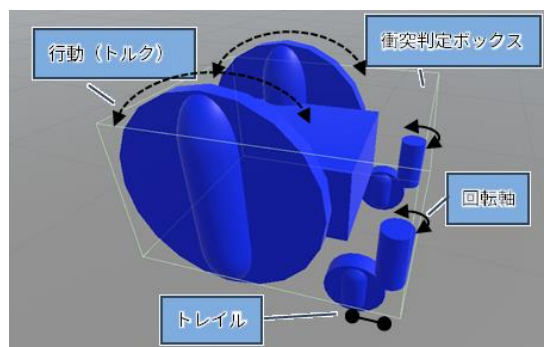


図1 エージェント（車椅子）

3.2. 住宅空間

検証に用いた住宅空間について説明する。今回は図2に示すような住宅内部の一角で学習を進める。バリアフリー向けの住宅を設計する際、通常の住宅を設計することと比べ様々なことに配慮し、寸法を決定する必要がある。参考文献⁴⁾によれば、有効開口幅が 950mm以上とのことであるので、チェックポイント（後述）のある開口の幅はいずれも 1,050mm とした。

学習のスタート地点は、上述の衝突判定ボックスの違いにおける学習の推移に着目することを目的に、図2に示すようチェックポイント①を設置した開口に正対した地点としタスクの難易度を下げるようにした。予備実験にて、より複雑なタスクを学習させるテクニックであるカリキュラム学習なども試行したが、本報告では扱わないことにする。

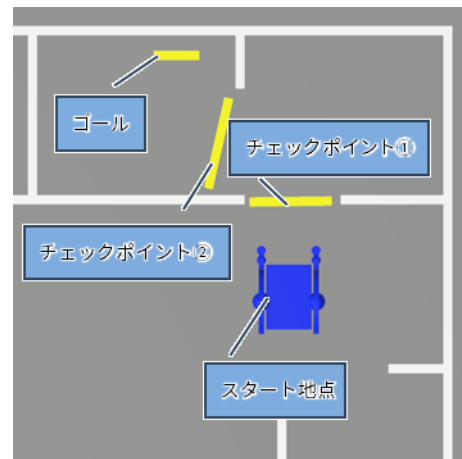


図2 住宅空間およびスタート地点とゴール

3.3. 報酬とエピソードの終了条件

図3を用い報酬とエピソードの終了条件について説明する。ゴールにはエージェントが衝突するとゴール判定となるオブジェクトを設置した。ゴール判定となると報酬（1.0点）を受け取りエピソードが終了する。他、壁に衝突してしまった場合、ゴールにたどり着かないまま 5000 ステップ経過した場合もエピソードが終了する。この場合、受け取る報酬は 0.0 点である。壁に衝突してもエピソードを継続することもできるが、この場合 5000 ステップ経過するまで壁への衝突をこともあるため、学習を非効率になる。今回は学習を効率的に行うため、壁に衝突したらエピソード終了とした。ところで、深層強化学習では、学習開始時にはエージェントにランダムな行動を行わせ、何らかの報酬を受け取らせることでポリシーが更新されていく。今回の環境のように、ランダムな行動のみではゴールまでたどり着けないことが予想される場合、その経路上に少量の報酬を受け取ることでできるチェックポイントを設置することが行われる。本研究でも2つの開口付近にチェックポイント①と②を設置する。エージェントがチェックポイン

トに辿り着いた際には 0.1 点の報酬を受け取り、エピソードが継続する。なお、各チェックポイントで報酬を受け取ることができるのはエピソード毎に一回のみである。

獲得報酬の平均値が 1.2 点（つまり 2 つのチェックポイントを通過して最終的にゴールに到達した状態）に近づくか、あるいはステップ数が 500 万回に達した場合、学習を終了する。

また、以上の直接的な報酬に加えて、Curiosity 報酬を加えた。Curiosity 報酬とは、エージェントが未知の状態に訪れたことに対して与えられる報酬であり、エージェントに好奇心を持たせ、未知の状態への探索を促すことを可能とする。エージェントに要求する動作が複雑であり、直接的な報酬を受け取ることのできるゴールやチェックポイントにたどり着くことが困難であると予想されたため、これを取り入れた。

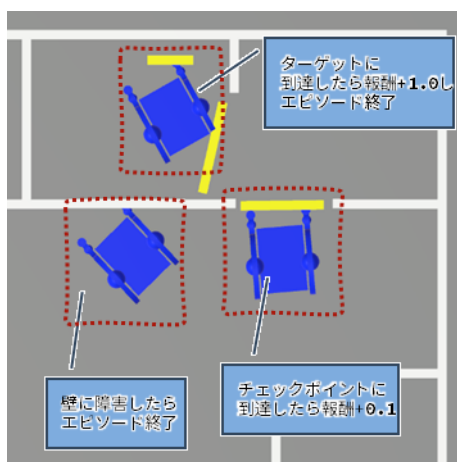


図3 報酬の設定

3.4. 観察

図2および図3で示したように住宅空間の一角を上から捉えた画像を入力する Visual Observation を採用した。画像の幅と高さは、両方 84 ピクセルに設定した。Visual Observation は観察対象の位置情報や速さ、向きなどを観察の情報とする Vector Observation とは異なり、カメラ情報のみを観察の情報とするため、壁などの障害物の認識に有効だと考えた。これを 128 ユニット、3 層のネットワークに入力し行動を求める。

なお、1 ステップ毎に観察をネットワークに入力し、行動を求めることもできるが、本研究では、3 ステップに一回この処理を行っている。

4. 検証結果

衝突判定ボックスの大きさを以下ように3パターンを設定し、間取りにおける開口などの有効寸法と、車椅子におけるバリアフリーの関係について検証した。さらに、パターン A' として、パターン A とボックスの大きさは同じだが、わずかに間取りを変えたパターンについても検証した。

4.1 ではパターン A に関して、4.2 では、衝突判定ボックスの大きさを変更したパターン B と C に関してまとめ、その結果について図4にまとめた。4.3 では間取りをわずかに変えたパターン A' に関してまとめ結果を図6に示す。

4.4 ではそれらのまとめてとして考察する。

- ・ パターン A : 800x500x900 (mm)
- ・ パターン B : 850x500x950 (mm)
- ・ パターン C : 900x500x1000 (mm)

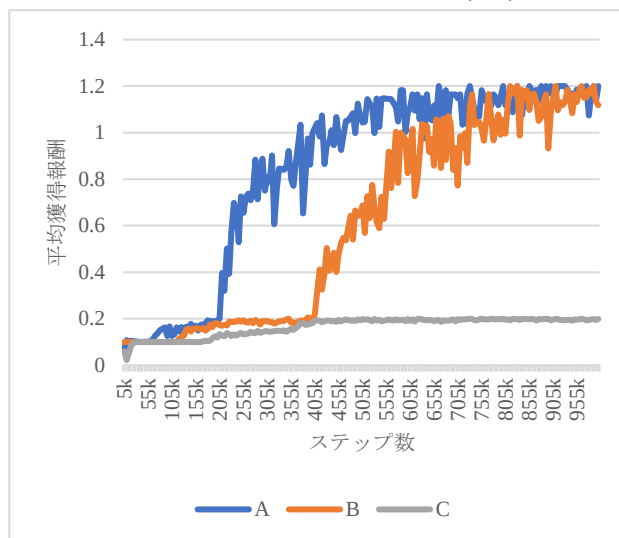


図4 パターンA,B,Cの平均獲得報酬の推移

4.1. パターン A

衝突判定用ボックスがパターン A はおおそ車椅子全体を取り囲む大きさであり、かつ自走する車椅子が要する幅とされる 0.8mを満たす。この大きさを最小の大きさとして検証する。

パターン A では約 20 万ステップを超えると、チェックポイント②に安定的にたどり着くようになり、以降時折ゴールにもたどり着くようになる。約 50 万ステップを超えると、90%の確率でゴールにたどり着くようになった。

図5はゴールにたどり着いた際の挙動の一例である。観察すると、エージェントはチェックポイント②に到達し、続いてゴールに向かう際に、直角に曲がる必要があるが、パターン A では滑らかに曲がる動作が見て取れた。この時の曲がり方は、片方（右）の駆動輪を前転させ、もう一方（左）の駆動輪にはトルクを加えない行動をとっており、結果、滑らかに曲がる挙動をしていた。自走式の車椅子では、片方を前転させ、他方を後転させることで方向転換して曲がる動作も可能である。エージェントを用意する際にはこの動作が行えるよう注意したが、パターン A の学習過程においてはこのような動作は見られなかった。前後するが、他のパターンにおいても、基本的には滑らかに曲がる動作のみが学習されることが多く、方向転換して曲がる動作の学習については、引き続き検討したいと考えている。

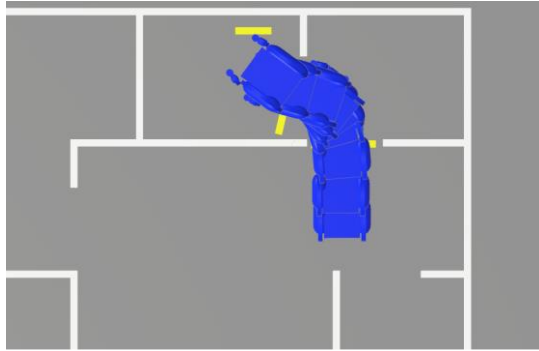


図5 パターン A がゴールにたどり着く様子

4.2. 衝突判定用ボックスの大きさを変更した場合

パターン B とパターン C

パターン B では、衝突判定用ボックスを 50 mm ずつ大きい 850x500x950mm とした。この場合、約 40 万ステップを超えると、チェックポイント②に安定的にたどり着くようになり、約 75 万ステップを超えると、90%の確率でゴールにたどり着くようになった。

パターン C ではさらに大きい、900x500x1000mm とした。この場合、チェックポイント②にたどり着くことができたものの、さらにそこから曲がりゴールにたどり着くことはできなかった。学習を 500 万ステップまで継続してみたがゴールには至らなかった。

4.3. 開口のデザインを変更した場合-パターン A'

間取りにおける開口などの有効寸法と車椅子におけるバリアフリーの関係が寸法にのみ依存するのかを検証するためにパターン A'を設定した。パターン A'は衝突判定ボックスの大きさはパターン A と同じであるが、チェックポイント②のある開口の位置を 15 mm 上にずらしたものである。学習の結果は図 6 のようになり、ゴールにたどり着くまでのステップ数はパターン A と比較して大幅に増加した。挙動を観察してみると、チェックポイント②に向かう際に、これまでのパターンでは稀であった方向転換によって曲がる動作が見られた。

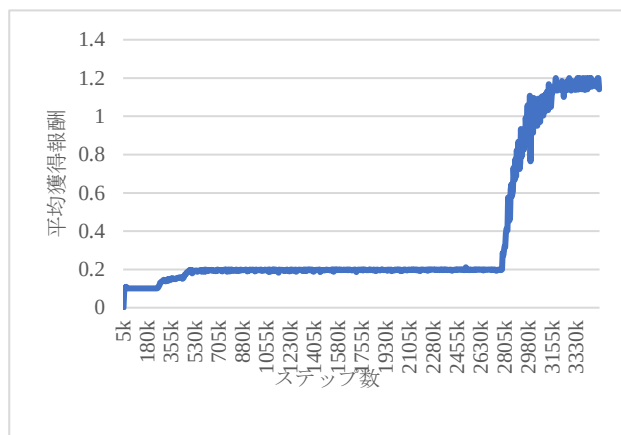


図6 パターン A'の平均獲得報酬の推移

4.4. 考察

パターン A の結果から、深層強化学習において、車椅子を操作しゴールにたどり着くことは実現できることが確認できた。主として曲がる際には滑らかに曲がる挙動を取る傾向が確認された。

パターン A～C を比較してみると、パターン A とパターン B はゴールにたどり着くことはできるが、それに要したステップ数には差があることが確認できる。今回のエージェントの学習における結果からいえば、この間取りの車椅子に対するバリアフリーの程度は、パターン A と B ではゴールにたどり着け、C ではたどり着けない程度、ということができる。

パターン A'の結果から、開口の有効寸法が同じであっても、学習の傾向が異なることが確認できた。深層強化学習により得た結果が実際の車椅子に十分に対応するかは議論の余地があるが、有効寸法のみでなく、開口付近の形状の違いも車椅子におけるバリアフリーに影響を与えることが示唆された。

5. まとめ

本研究では、車いすがバリアフリー住宅内部でゴールに行くまでの行動を仮想環境下で観察し、バリアフリーの有効寸法と衝突判定ボックスの関係について検証し、本研究で用いた手法がバリアフリー度合の評価に有効であることを示した。しかし、実際の車椅子の挙動に見られる直角に曲がるために片方の車輪を前転させ、他方を後転させ直角に曲がる動きが見られにくいことや初期条件の違いによって学習に差が生じることなど課題が浮き彫りになった。今後はこれらを精査し、バリアフリーの評価法について再度検討していく所存である。

【注釈】

- 注1. Unity Technologies の開発するゲーム向け統合開発環境。建築分野においては VR 開発にも用いられる。
- 注2. Unity において、深層強化学習を行うためのフレームワーク。
<https://github.com/Unity-Technologies/ml-agents> (2023-09-29 アクセス)

【参考文献】

- 1) 藤澤、他、深層強化学習を用いた自律搬送機械の運搬作業の動作生成に関する研究-実在建築物を用いた環境での運搬シミュレーション-、日本建築学会大会学術講演梗概集（北海道）材料施工、pp.199-200、2022
- 2) 藤岡、他、深層強化学習による五軸加工機の迂回パス生成に関する研究、日本建築学会技術報告集、27 巻 67 号、pp.1553-1558、2021
- 3) 高添、他、高添、他、BIM と強化学習を用いたドローンナビゲーションに関する基礎的研究、日本建築学会大会学術講演梗概集（近畿）情報システム技術、pp.125-126、2023
- 4) 「住まいの建築計画」、今村仁美、田中美都、学芸出版社、2021