

卒業論文 2019 年度 (令和 1 年)

目的満足度に応じたモビリティ全体制御

慶應義塾大学環境情報学部環境情報学科
島津翔太

目的満足度に応じたモビリティ全体制御

自動運転技術の発展により、我々人間は車を常時コントロールする必要がなくなってきた。例えば、車線の自動変更や物体検知による自動ブレーキなどは、今まで人間が行っていた操作をコンピューターによって行っている。今後、自動運転技術はより高度になり、近い将来人間がハンドルを握らなくともコンピューターによる制御のみで目的地まで到達できるレベルまで発展する可能性がある。そのような自動運転車が普及した社会において、自動車各個が個別に行動すると様々な問題が発生する。

第一に、特定の経路の混雑があげられる。多数の自動運転車が個別に経路を選択した場合、特定の道が混雑する問題が発生する。第二に、自動運転による使用用途の変化への順応である。人間による操作が一切行われない完全自動運転が実現する事はすなわち無人運転が可能であることを意味する。無人運転が可能となった場合、自動車の保有者が目的地についた後に駐車する必要はなく、駐車時間中に自動車を使いたい人の元へ迎えに行く無人のヒッチハイクのような行為が可能となる。このような使用用途の変化に対応するには、単に目的地までのルートを選択するだけではなく目的地に到達した後の行動も自動運転車が決定する必要がある。

そこで、本研究では強化学習を用いて自動車の利用者の目的に合わせた経路選択の実験を行う。実験では、緊急のミーティングに向かっている、空腹を感じているなどの利用者の目的や状態などのパターンを予め想定した。そのパターンに対して利用者の目的を達成するルートを選択した場合に強化学習に報酬値を与え、目的を満たせないルートを選択した場合は罰則値を与える。これを繰り返すことにより、強化学習アルゴリズムがより目的を達成するような経路選択を行えるようになると考えられる。

これらの実験を通して、本研究では正解データなし自律して意思決定を行うアルゴリズムの一種である強化学習を用いたアプローチが人間の目的を満たす意思決定に有効であること。モビリティの経路選択において実用的であることを示す事を期待する。

キーワード:

1. 機械学習, 2. 強化学習, 3. モビリティ, 4. シェアリングエコノミー

慶應義塾大学環境情報学部環境情報学科
島津翔太

Mobility for Human Satisfaction

With the development of autonomous driving technology, humans no longer need to constantly control cars. For example, automatic lane changes and automatic braking based on object detection are done by computers in the same way that humans have done so far. In the future, automatic driving technology will become more advanced, and there is a possibility that in the near future, it will be developed to a level where a person can reach a destination only by computer control without grasping the steering wheel. In a society where such self-driving vehicles are widespread, various problems arise when individual vehicles act individually.

First, there is congestion on specific routes. If a large number of autonomous vehicles select their own routes individually, there will be a problem of congestion on specific roads. The second is to adapt to changes in usage by automatic operation. The realization of fully automatic operation without any human intervention means that unmanned operation is possible. When unmanned driving becomes possible, it is not necessary to park a car after the car owner arrives at the destination, and an act such as an unmanned hitchhiking to pick up a person who wants to use the car during parking time becomes possible. In order to respond to such changes in usage, it is necessary for the self-driving vehicle not only to select the route to the destination but also to determine the actions to take after reaching the destination.

In this study, therefore, we use reinforcement learning to experiment with route selection according to the purpose of vehicle users. In the experiment, we assumed the user's purpose and condition such as going to an emergency meeting or feeling hungry. If the user selects a route that achieves the user's purpose for that pattern, we give a reward value to reinforcement learning. If the user selects a route that does not meet the user's purpose, we give a penalty value. By repeating this process, the reinforcement learning algorithm may be able to select a route that achieves a better goal.

Through these experiments, we hope to show that an approach using reinforcement learning, which is a kind of algorithm that makes independent decisions without correct data, is effective for making decisions that satisfy human objectives and is practical for path selection of mobility.

Keywords :

1. Machine Learning, 2. Reinforcement learning, 3. Mobility, 4. Sharing Economy

Keio University Faculty of Environment and Information Studies
Shota Shimazu

目次

第1章	序論	1
1.1	本研究の目的	1
1.2	本論文の構成	1
第2章	背景	2
2.0.1	自動運転	2
2.0.2	Mobility as a Service	2
2.0.3	シェアリングエコノミー	3
2.0.4	機械学習	3
第3章	本研究における問題定義と仮説	4
3.1	最短経路問題	4
3.2	本研究における問題定義	5
3.2.1	強化学習によるルート選択	5
3.3	仮説	5
第4章	提案手法	6
4.1	概要	6
第5章	要素技術	7
5.1	機械学習	7
5.2	深層学習	7
5.3	深層強化学習	7
5.4	Deep Q Neural Network	8
第6章	実験	9
6.1	概要	9
6.2	構成	9
6.3	Deep Q Neural Network Model	10
6.4	Data Serializer	10
6.5	可視化地図アプリケーション	10
第7章	評価	11
7.1	評価内容	11

第 8 章	結論	12
8.1	本研究のまとめ	12
8.2	本研究の課題	12
付 録 A	APPENDIX	13
A.1	本論文で用いる機械学習関連の略語	13
謝辞		14

図 目 次

3.1	最短経路問題	4
3.2	強化学習による経路検索の結果	5
5.1	強化学習による経路検索の結果	8
6.1	システムの関係図	9

表 目 次

第1章 序論

本章では本研究の動機，本研究の概要を示す．

1.1 本研究の目的

本研究では，強化学習を用いた経路選択を試行し自動車の利用者の移動目的をより多く達成できる強化学習モデルの構築を目指す．現状，自動運転車に代表されるモビリティは個別に行動しており，ルートの選択に置いても人間が与えた目的地への最短経路を選択している．この場合，一定時間内に同一方面の目的地を多数設定された場合に特定のルートの混雑が予測される．また，レベル5の完全自動運転が実現された場合，自動車を使用していない期間は自動で他者に貸し出すなどの利用方法の変化も見込まれる．

本研究では，機械学習モデルの一つである深層強化学習を用いて，経路選択の実験を行い，自律意思決定型の機械学習アルゴリズムが自動運転車において活用可能であることを示したい．

1.2 本論文の構成

本論文における以降の構成は次の通りである．

2章では，背景を述べる．3章では，本研究における問題の定義と，解決するための要件の整理を行う．4章では，本研究の提案手法を述べる．6章では，4章で述べたシステムの実装について述べる．7章では，3章で求められた課題に対しての評価を行い，考察する．8章では，本研究のまとめと今後の課題についてまとめる．

第2章 背景

本章では、本研究の背景を示す。

2.0.1 自動運転

車における自動運転は1980年代から研究されてきた。例えば、欧州で1987年から1995年に行われたEUREKAプロメテウス計画では高速道路における車線の追従や車線の変更などの自動運転の基礎技術が研究された。現在では、これらの機能は市販の自家用車にも運転を支援する機能として搭載されている。また、高速道路など限定した場所であれば人間による介入が不要な一部自動運転が可能となっているものもある。今後、将来自動運転技術はより人間による操作を少なくし、首相官邸ホームページ「官民 ITS 構想・ロードマップ 2019 [1]」に定義されたレベル5の完全な自動運転技術も完成する可能性がある。

2.0.2 Mobility as a Service

日本に置いて、車や鉄道などの交通は高度経済成長期以降、道路網や路線網の拡大も合わせて急速に普及が進み、旅客・貨物共に主たる移動手段となった。

しかし、近年、交通は単なる移動手段としてだけではなく、移動や移動に付随する付加価値や自己所有の車を自ら運転すると行った従来の使い方からの変化が求められるようになってきた。

これに対して移動をサービスとして提供するという考え方があり、Mobility as a Service (通称: MaaS) と呼ばれている。現在 MaaS サービスとしては米 Uber などに代表される個人所有の車を配車するサービスや、自動車を不特定多数の利用者で共有するカーシェアリングサービスなどがある。

人間による操作を必要としないレベル5の自動運転が実現すると、ハンドルを握る必要がないため従来のように自動車において移動時間中に運転に拘束されることがない。また、移動経路も人間が考えることなく目的地まで到達する。このようになると、単に自動車そのものを共有するサービスだけではなく移動経路の選択や移動時間の活用などのサービスとして提供する事が必要になると予想される。

2.0.3 シェアリングエコノミー

シェアリングエコノミー [2] ないしは共有経済とはモノやサービスを特定の個人で所有するのではなく複数人で共有する社会関係である。古くは GNU プロジェクトなどのオープンソースソフトウェアがそれに当たると考えられている。スマートフォンの普及により、Uber などの個人所有の車の配車サービスや AirBnb などの所有する不動産を一定期間旅館のような形で貸し出すと行ったサービスが登場した。将来、人間が介在することのない自動運転が可能になると自動車そのもののハードウェアをシェアするだけではなく自動車を使ったサービスのシェアが加速すると思われる。

2.0.4 機械学習

機械学習とはコンピューターが自動的にパターンを学習し人間による指示がなくとも課題を実行する事である。モバイルコンピューターの性能が飛躍的に向上し画像の認識や機械翻訳などの多くの分野で応用されるようになってきた。これらの技術は日々進化を重ねており画像分類の分野においては 1 年で認識率を 20% 向上したり、学習にかかる時間を 50% ほど短縮するアルゴリズムが考案されたりと飛躍的に向上してる。また、機械学習アルゴリズムは意思決定を伴う処理にも活用され、AlphaGo に代表されるように囲碁などのボードゲーム分野において人間をも上回る成果を出している。加えて、意思決定型の機械学習はボードゲーム意外にも応用されつつあり、リスティング広告の最適化など利用者である人間にとっての利益が最大化するような制御に応用されつつある。

第3章 本研究における問題定義と仮説

本章では、第2章で述べた背景から、現状の自動運転システムが目的としている制御方法の問題点を述べる。

3.1 最短経路問題

自動運転が行われるようになると、人間による非合理的なブレーキなどの操作がなくなり渋滞が解消すると言われている。しかし、渋滞は車線あたりの交通量に比例し、一般的に自動車の走行ルートを決定的する場合、現在地から目的地までの最短ルートを選択するため交通需要の高い経路の交通量は変わらず、自動運転による渋滞解消の因果関係には疑問が残る。少なくとも同時時間帯に同一地点付近の目的地を設定した車が大量にいた場合に特定のルートが混雑が発生することは避けられないと考えられる。現状のカーナビゲーションシステムなどに搭載されているような渋滞回避機能であっても、最短経路の周辺の道路に回避され最短経路を選択するには早いもの順になる状態になる。



図 3.1: 最短経路問題

3.2 本研究における問題定義

現状の強化学習の問題点を列挙し, 整理する.

3.2.1 強化学習によるルート選択

現状強化学習を経路選択に応用した事例 [3] はすでに存在する. 例えば, 下図のような単純な正方形の升目上の任意の座標にスタート地点とゴール地点を設定し, 最短経路を学習するという物である. しかし, 実際の道路網は迷路と比較しても非常に複雑であり沿線との接続や経由地などの要素を考慮す強化学習を応用した場合の精度は未知数である.

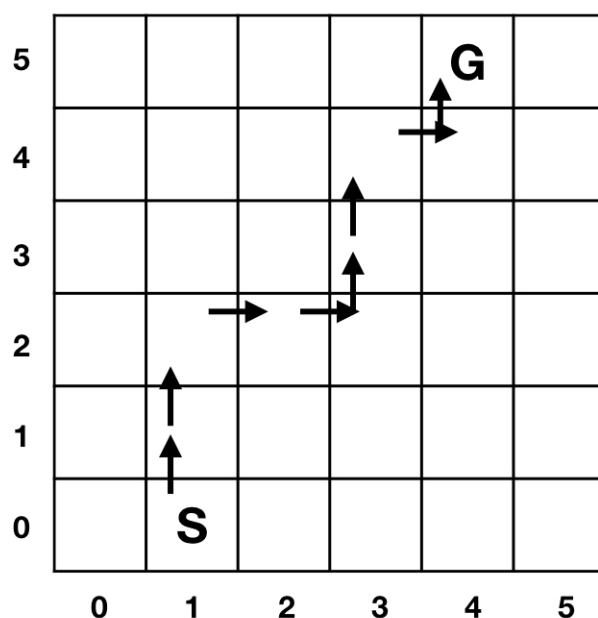


図 3.2: 強化学習による経路検索の結果

3.3 仮説

Deep Q Nueral Network により学習を繰り返すことで, シミュレーターないでないので利用者に最も最適化された経路選択を行うようになると考えた.

第4章 提案手法

本章では提案手法について述べる.

4.1 概要

本研究では, 人間の状態とモビリティの経路を再現するシミュレータープログラムを開発し実験を行う. なお, シミュレータープログラムの作成にあたっては以下について仮定を行った.

- シミュレーター上での自動車は全てレベル5の完全自動運転ができる物とする
- 歩行者や, 自動運転非対応者などは考慮していない
- シミュレーターでは予め決められた幹線道路及, 環状線などのバイパス道路のみを考慮し, それ以外の道に関しては考慮しない
- 人間の目的を正確に推定するコンピューターシステムもしくはセンサーなどは実在しているものとする.

第5章 要素技術

本章では要素技術について述べる

5.1 機械学習

機械学習とは広義には、コンピューターが自動的にパターンを学習し人間による明示的な命令がなくとも特定の課題を自動で実行する技術又はアルゴリズムのことである。主に、正解データを与えることによってパターンを学習する教師あり学習、データのまとまりや相関を求める教師なし学習と強化学習に分類される。

5.2 深層学習

深層学習とは脳が持つ脳神経系をソフトウェアで再現した人工ニューラルネット (ANN) を持つ機械学習アルゴリズムの一つである。人間の脳を模したパーセプトロンによる深層学習自体は 1957 年から提唱されていた。4 層以上のパーセプトロンでは過学習や勾配消失問題が発生しコンピューテーションコストも大きいためあまり普及しなかった。

5.3 深層強化学習

深層強化学習とは、強化学習に深層学習を組み合わせた機械学習アルゴリズムである。強化学習はエージェントとエージェントが動作する環境を定義し、定義された環境下でエージェントへの報酬が最大化するように学習を行う。

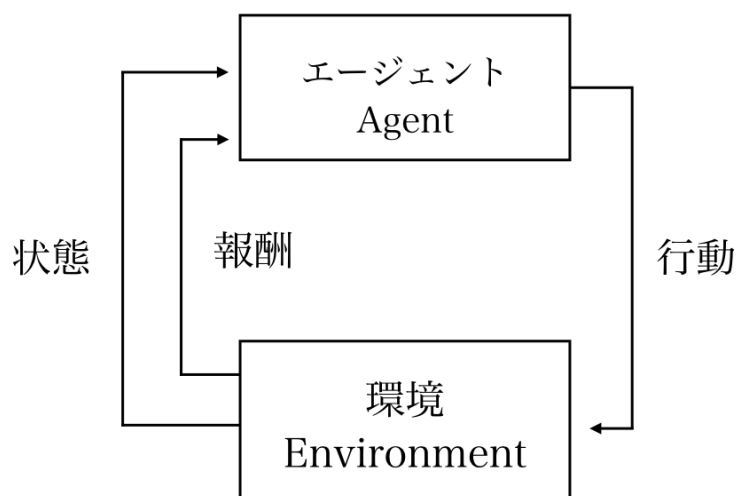


図 5.1: 強化学習による経路検索の結果

5.4 Deep Q Neural Network

Deep Q Neural Network(DQN [4]) は Q 学習という強化学習における古典的なアルゴリズムを深層強化学習に応用したものである.

第6章 実験

本章では提案手法の実験について述べる.

6.1 概要

本研究では, 以下に述べるコンピュータプログラムを実装し, モビリティや人の動きをコンピュータ上で再現する. 作成するプログラムは 第一は, 強化学習モデルを生成する機械学習プログラムと

6.2 構成

この実験での各プログラムやミドルウェアの関係性をいかに示す.

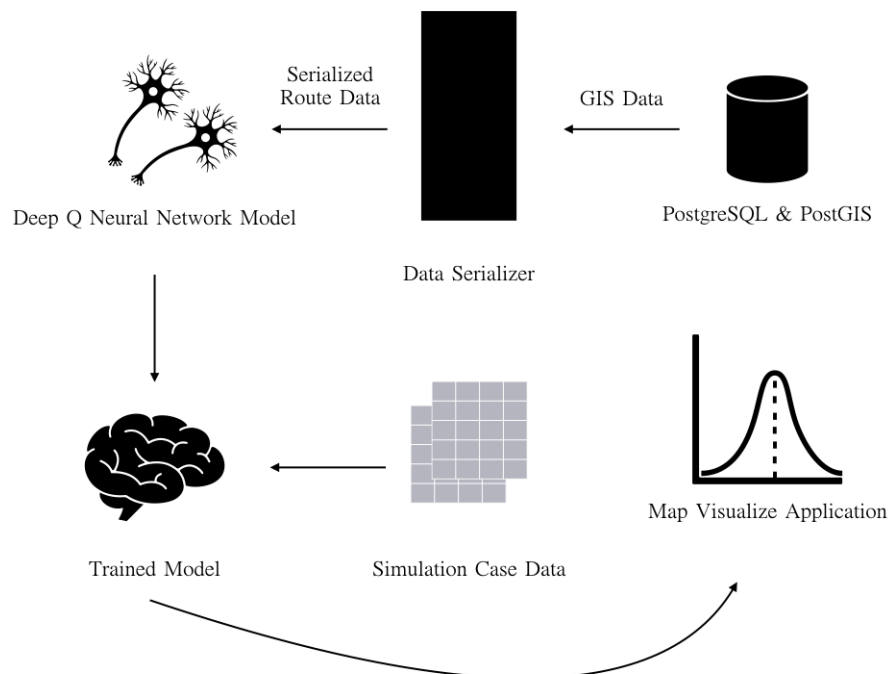


図 6.1: システムの関係図

6.3 Deep Q Neural Network Model

このプログラムは Deep Q Neural Network による強化学習を行うプログラムである。実装言語は Python である。

6.4 Data Serializer

Data Serializer は PostgreSQL サーバー上に記録された経路などの地理データを上述した Deep Q Neural Network に学習させるためのデータ変換を行うプログラムである。開発言語は Python であり、一部データベースとの接続部分に C++ を用いた。

6.5 可視化地図アプリケーション

モビリティの動きを地図上に可視化するアプリケーションの作成を行った。このアプリケーションは Deep Q Neural Network の学習済モデルが下した意思決定つまり経路を地図上にアイコンとして再現し、乗車している人間に関する情報を統合して閲覧できるようにしたものである。

ライブラリ名 / 言語など	バージョン	実行デバイス
Swift	5.1.3 (swiftlang-1100.0.282.1 clang-1100.0.33.15)	iPad Pro
React Native	0.61.5	iPad Pro

第7章 評価

本章では，提案システムの評価について述べる．

7.1 評価内容

本研究の実験において以下の項目について評価を行う．

- 人間の状態すなわち目的をレベル別に何個満たせたか
- 1km メッシュにおける人口の変異を時系列方向に観測し，全体的に移動が起きているか

第8章 結論

本章では，本研究のまとめと今後の課題を示す．

8.1 本研究のまとめ

実験が全て完了したのちにまとめる

8.2 本研究の課題

本綴じに向けて真剣に取り組む

付 録 A APPENDIX

A.1 本論文で用いる機械学習関連の略語

- ANN - Artificial Neural Network
- DNN - Deep Neural Network
- CNN - Convolutional Neural Network
- など

謝辞

俺に関わった全てに感謝.

参考文献

- [1] 官民データ活用推進戦略会議 高度情報通信ネットワーク社会推進戦略本部. 官民 ITS 構想・ロードマップ 2019. <https://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20190607/siryou9.pdf>, 2019.
- [2] Lacy Peter and Rutqvist Jakob. サーキュラー・エコノミーで競争優位性を確立する. https://www.accenture.com/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Local/ja-jp/PDF_4/Accenture-Waste-Wealth-Exec-Sum-JP.pdf, 2019.
- [3] Shibata Katsunari. Learning of exploration behavior by reinforcement learning. <http://shws.cc.oita-u.ac.jp/~shibata/pub/SSI05.pdf>, 2008.
- [4] Mnih Volodymyr, Kavukcuoglu Koray, Silver David, Rusu Andrei A., Veness Joel, Bellemare Marc G., Graves Alex, Riedmiller Martin, Fidjell Andreas K., Ostrovski Georg, Petersen Stig, Beattie Charles, Sadik Amir, Antonoglou Ioannis, King Helen, Kumaran Dharshan, Wierstra Daan, Legg Shane, and Hassabis Demis. Human-level control through deep reinforcement learning. <https://www.nature.com/articles/nature14236>, 2015.