

仮想環境下における 深層強化学習を用いた自動運転車での ラウンドアバウト走行検証

日本大学 文理学部 情報科学科 谷聖一研究室

南 基大

2020/1/26

目次

1. はじめに

- 1. 演習動機
- 2. 演習目的
- 3. 演習内容

2. 準備

- 1. 強化学習
- 2. 深層強化学習

3. 演習

I. 実装

- A) 実装環境
- B) 訓練用環境: サーキット
- C) 評価用環境
- D) エージェント: 自動運転車

II. 学習

- A) 学習手順
- B) ニューラルネットワーク
- C) 報酬
- D) 学習順組み合わせ
- E) 学習の観測

III. 評価

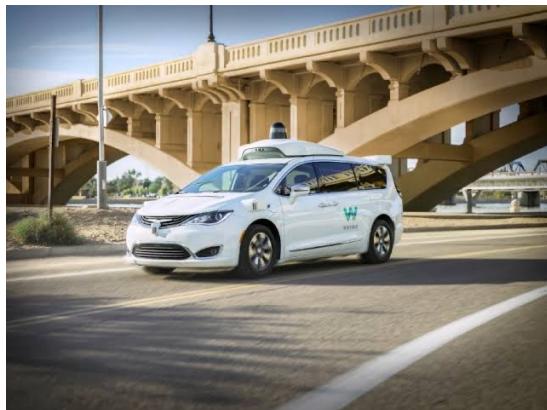
- A) 評価基準
- B) 結果
- C) 考察

4. 今後の課題

1-1 演習動機

2018年エンジニアの聖地「シリコンバレー」へ半年間留学中

Waymo社の自動運転車両が目の前を通り



<https://waymo.com/>



1-1 演習動機

「自動運転なんてまだ未来の話でしょ？」
と思っていたら

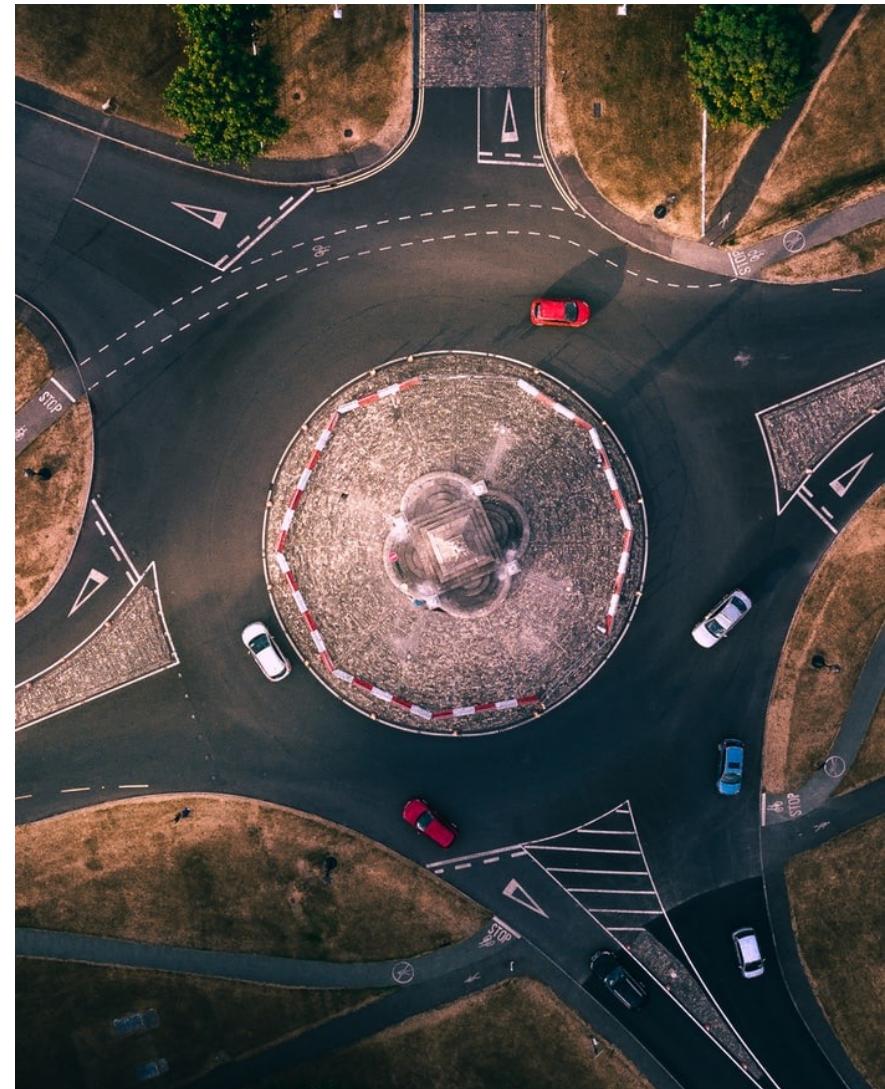
夢の自動運転車が目の前を走る衝撃

自動運転技術へ興味を持つ

1-1 演習動機

欧米滯在中, ラウンドアバウトに遭遇
戸惑うが直感的に走行出来た

自動運転だったら?



1-2 演習背景（自動運転）

自動運転

カメラやレーダー、GPS等で周囲の環境を認識し、
人工知能によって**運転手の補助**もしくは**自律運転**を行う

米国自動車技術協会によって
自動運転レベルが5段階で定義

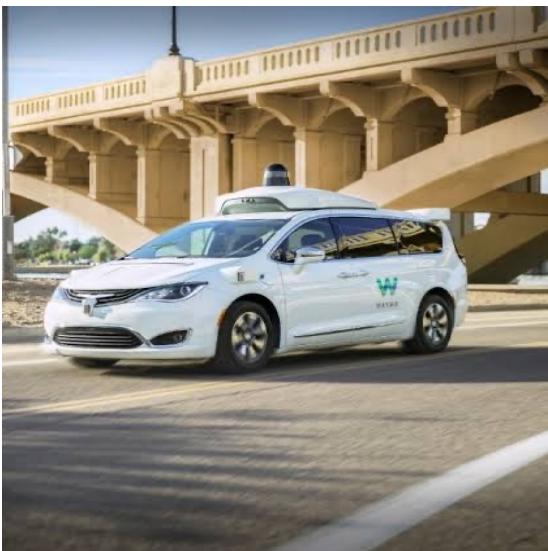
レベル4以上で操縦不要

(米国自動車技術会=SAEの定義)

人が主体	レベル0	自動運転機能のない一般的な車
車が主体	レベル1	自動ブレーキなどの運転支援機能
車が主体	レベル2	部分的な自動運転。運転手は常に監督する必要
車が主体	レベル3	条件付きの自動運転。緊急時をのぞき運転を車に任せる
車が主体	レベル4	高度な自動運転。環境は限られるが運転手の対応は不要
車が主体	レベル5	完全な自動運転。運転手も不要

1-2 演習背景（自動運転）

様々な企業がレベル4(無人)自動運転を実現



<https://waymo.com/>



<https://global.toyota/jp/>

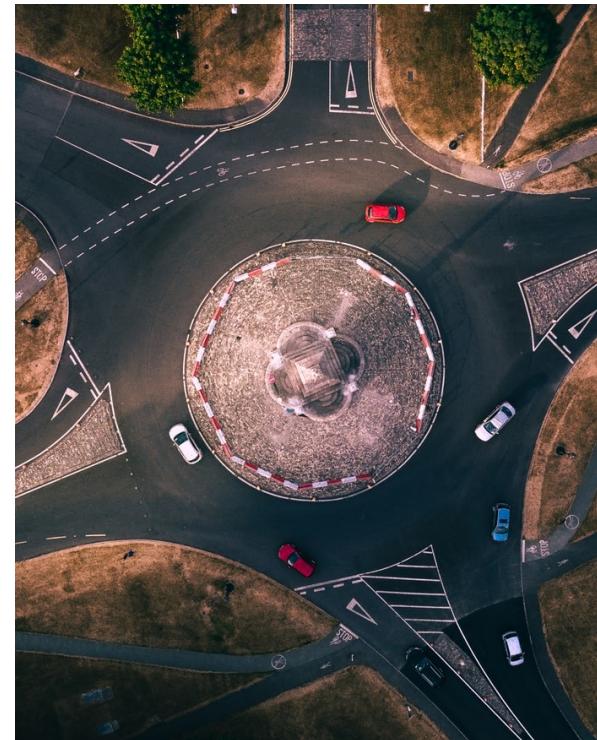


<https://tier4.jp/>

1-2 演習背景 (ラウンドアバウト)

特徴

- ロータリー(円形交差点)の一種
- それぞれの道路が環状道路によって接続



1-2 演習背景（ラウンドアバウト）

先行研究にて事故発生率及び燃費等に関する
ラウンドアバウト(環状交差点)の手動運転への統計的優位性が示されている

The findings of these studies all show that **injury crashes are reduced** more dramatically than crashes involving property damage only.

"Roundabouts: An Informational Guide Second Edition", NCHRP Report 672 p136,FHWA,2010.

“Taking into account the above situation, as a conclusion, it should be emphasized that **reducing energy consumption**”

Hang Cao, Máté Zöldy, “An Investigation of Autonomous Vehicle Roundabout Situation”, 2019

1-2 演習背景（ラウンドアバウト）

現時点では自動運転でのラウンドアバウト通行は困難

As such, the AI either struggles to make it through the traffic circle safely, or at the last moment is **gives up and tosses control of the vehicle to the human driver.**

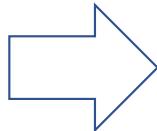
Solving the Roundabout Traversal Problem for Self-Driving Cars

<https://www.aitrends.com/ai-insider/solving-roundabouts-traffic-circle-traversal-problem-self-driving-cars/>

1-3 演習目的

手動運転では優れている

ならば



自動運転でも適用可能

かもしれない



深層強化学習を用いた自動運転車を作成
ラウンドアバウトへの適応を検証

1-4 演習内容

実験手順は以下の通り

交通システムと自動運転車を仮想環境に構築



自動運転車を深層強化学習を用いて学習



生成した自動運転車を用いて仮想環境にて検証

2-1 強化学習

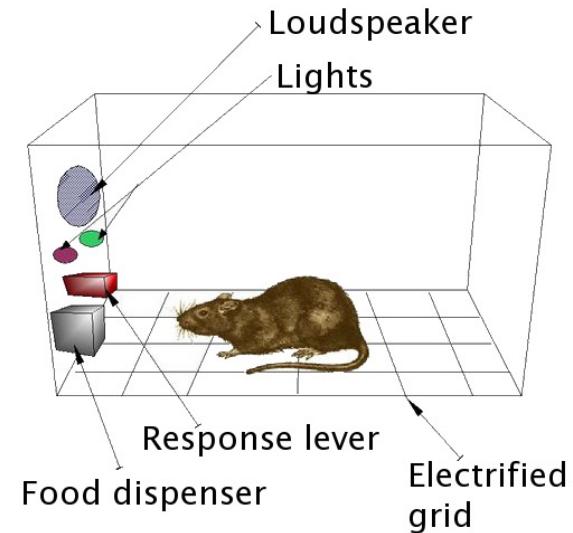
スキナーの箱(報酬学習)

レバーを押すことでエサが出てくる飼育箱にラットを入れる

偶然ラットがレバーを押して餌が出てくる

レバーを押すと餌が出ることを学習

ラットはより多くの餌を得る



Andreas1 - Adapted from Image:Boite skinner.jpg, CC 表示-継承 3.0,
<https://commons.wikimedia.org/w/index.php?curid=5709267>による

より多くの「報酬」 = 「餌」を得られるように試行錯誤、主体者の行動選択が変化

強化学習

2-1 強化学習

機械学習における強化学習の要素

エージェント (Agent)

観測し行動する主体者を表す

状態 (State)

ある時刻に環境がどうなっているかを表す

行動 (Action)

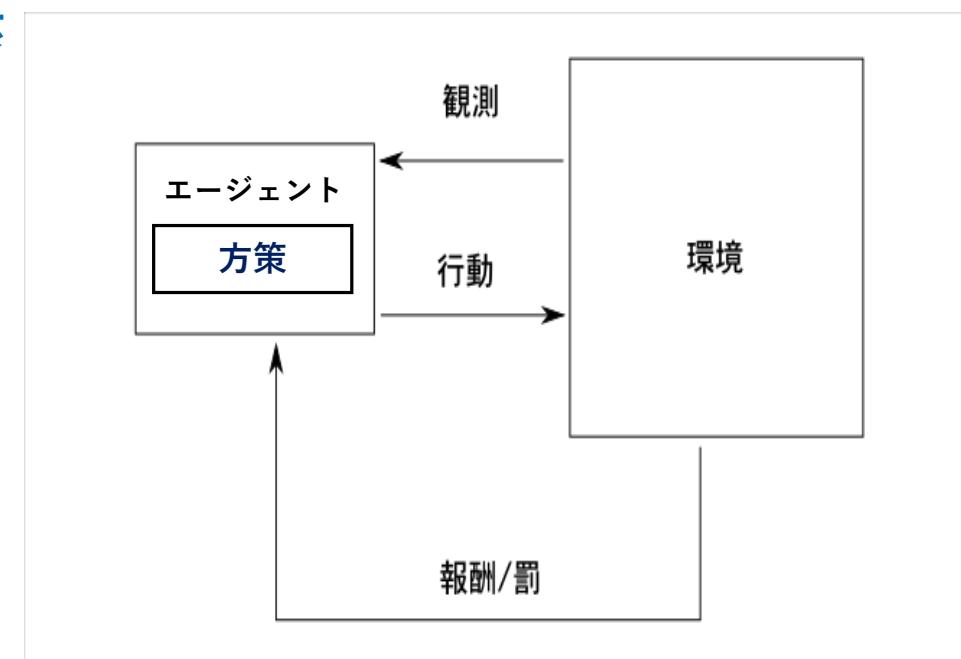
ある環境に対し,エージェントが取れる行動

報酬 (Reward)

ある状態に対し,エージェントがある行動を起こした際に
どれくらいの報酬を得られるかを表す

方策 (Policy)

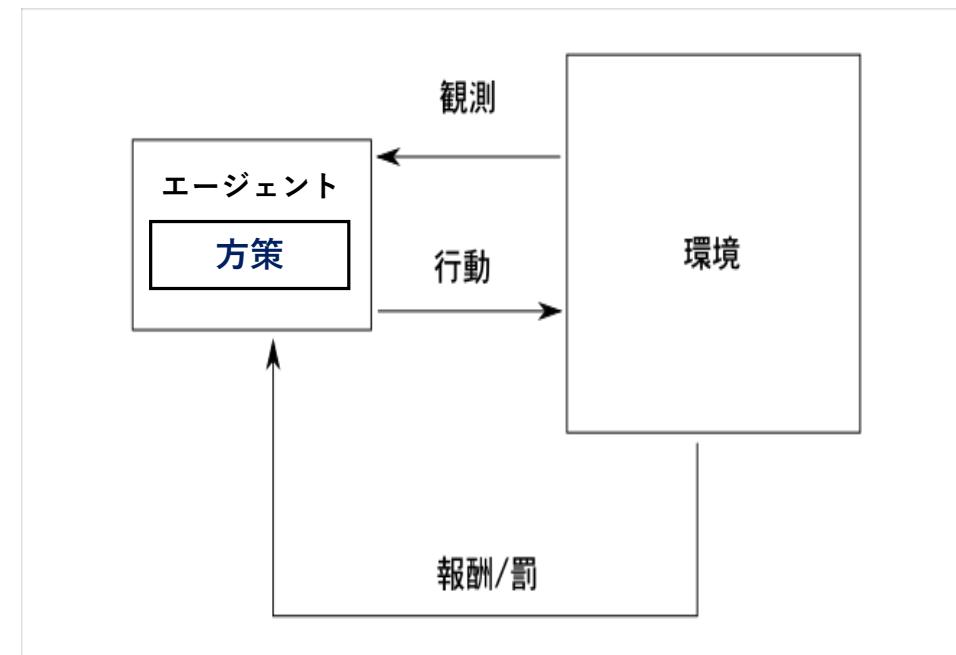
ある状態に対し,エージェントがどのような行動を
選択するかを決める



2-1 強化学習

機械学習における強化学習の特徴

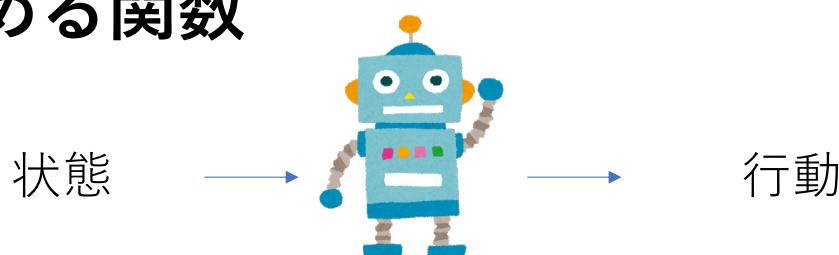
- ・報酬は開発者が設定
- ・問題の規模・種類によっては現実では試行錯誤が困難
→仮想環境で実行



2-2 深層強化学習

方策(Policy)

ある状態に対し,エージェントがどのような行動を選択するか決める関数

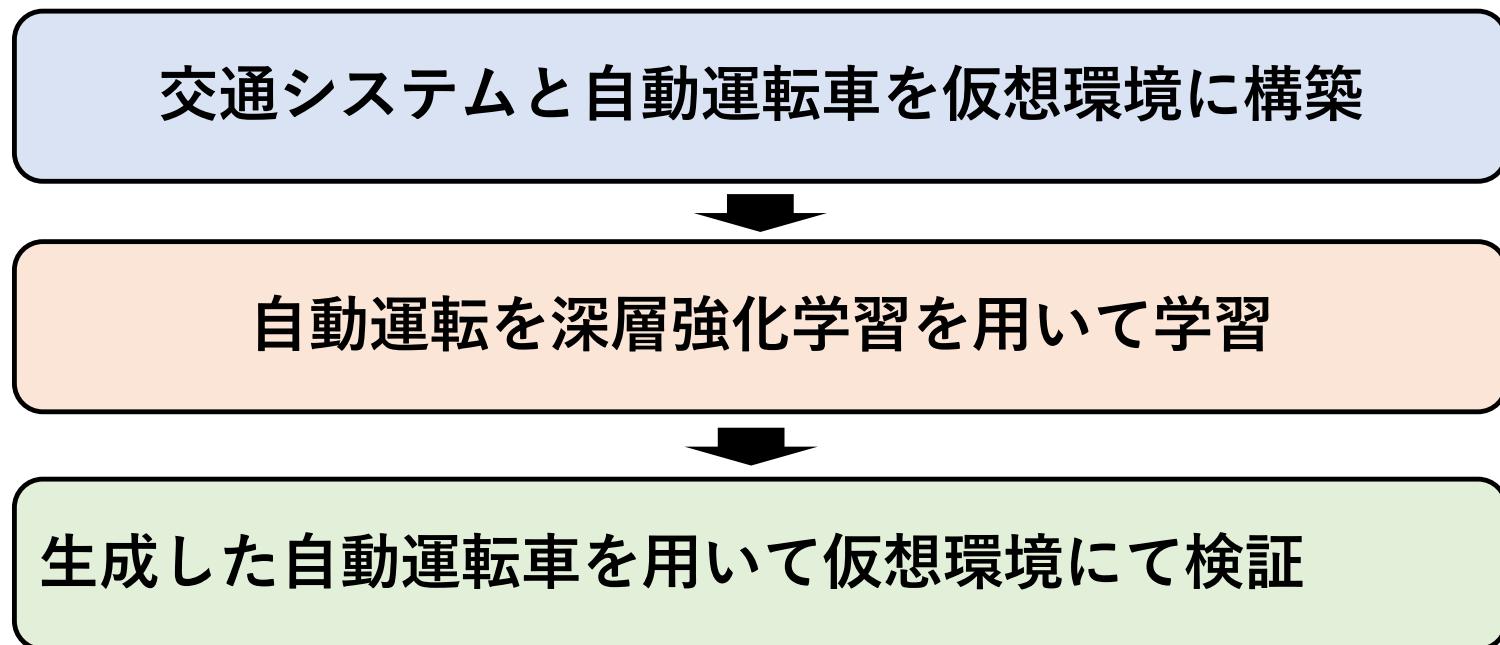


方策関数に深層ニューラルネットワークを利用
深層強化学習

3 演習

3 演習内容（再掲）

実験手順は以下の通り



3-I 実装

交通システムと自動運転車を仮想環境に構築



自動運転を深層強化学習を用いて学習



生成した自動運転車を用いて仮想環境にて検証

3-I-A 開発環境

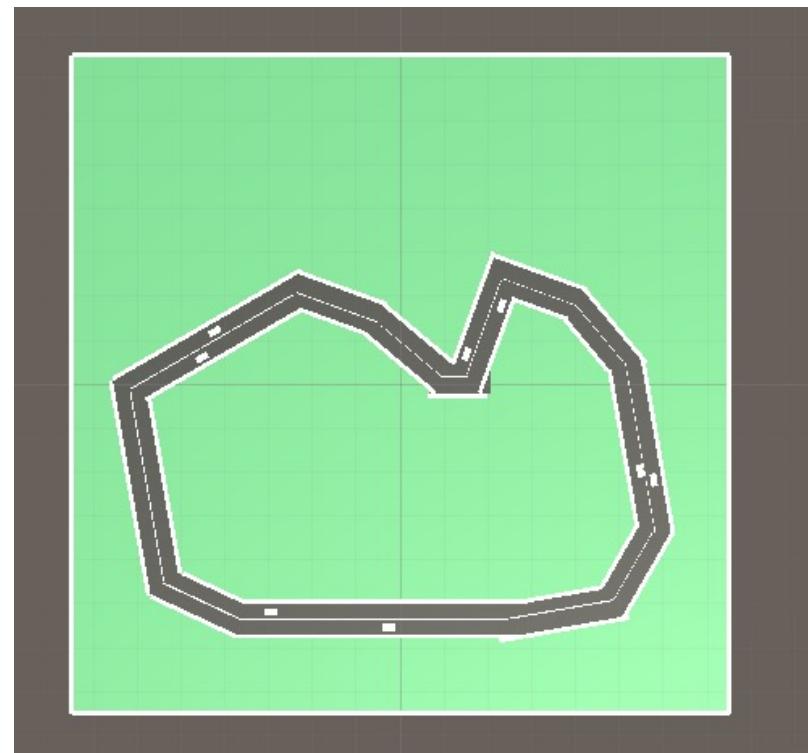
シミュレーション環境として統合開発環境の「Unity」を使用
環境の構築が容易



ゲーム向け深層強化学習ライブラリ「ML-Agents」
エージェントの設計が容易

3-I-B 訓練用環境: サーキット

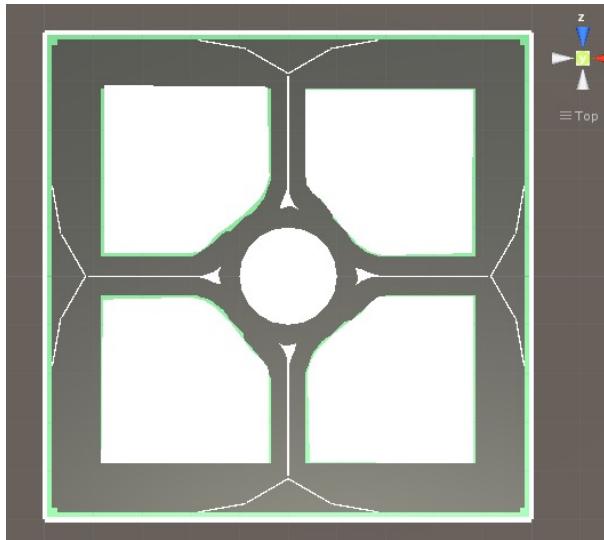
MAPサイズ :150m × 150m
道幅 :7m



3-I-C 評価用環境

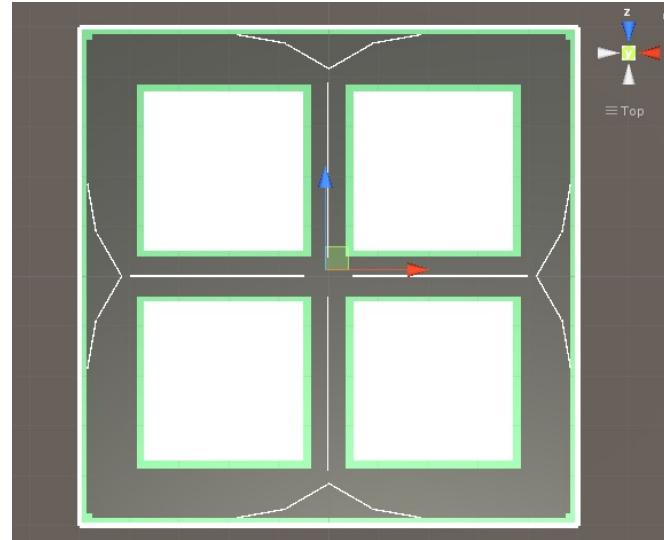
ラウンドアバウト

MAPサイズ : $100m \times 100m$
道幅
:中央十字路 7m
:環状道路 8-9m
:外周道路 9m



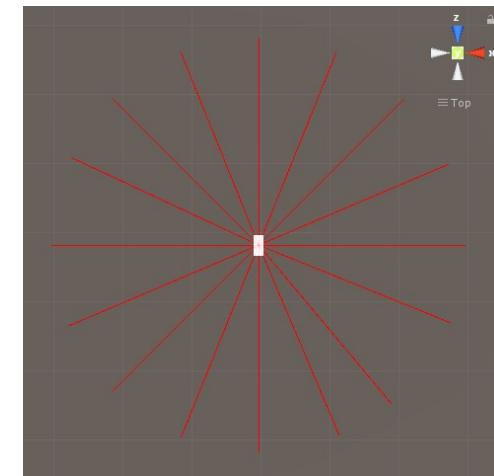
十字交差点

MAPサイズ : $100m \times 100m$
道幅
:中央十字路 7m
:外周道路 9m



3-I-D エージェント:自動運転車

- 自動運転車
(5,2)の長方形
車体中央にセンサーの中心部



センサー情報	車体情報	センサーを可視光線化したもの
センサー数	最高速度	60m /s
採取対象	旋回角度	3°
採取情報	加速量	20m /s ²

3-I-D エージェント:自動運転車

- 行動



Step毎に速度・ハンドルから各1つの行動を実行

3-II 学習

交通システムと自動運転車を仮想環境に構築



自動運転を深層強化学習を用いて学習



生成した自動運転車を用いて仮想環境にて検証

3-III-A 学習手順

訓練用環境にて一定回数まで学習
基本運転を習得



後述する組み合わせにて学習
交差点における運転を習得

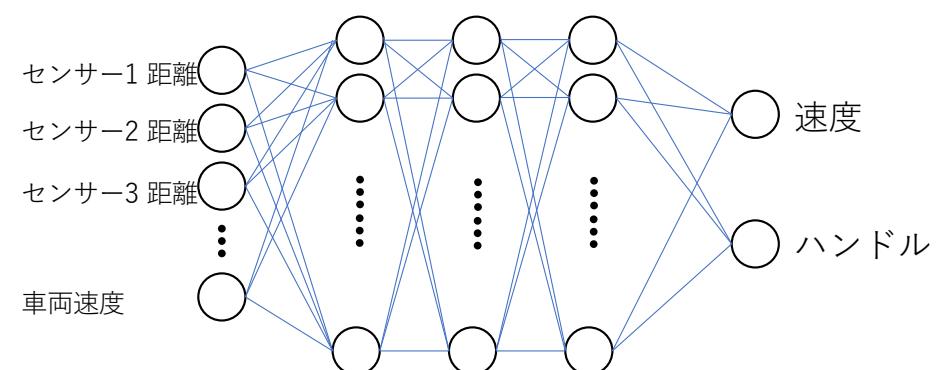
3-II-B ニューラルネットワーク

今事前実験で成績の良かった以下の値を使用

中間層数	活性化関数	学習率
3	ReLU	0.003

ノード数

入力層	各中間層	出力層
105	524	2



3-II-C 報酬

事前実験で成績の良かった以下の値を使用
各step毎に報酬を計算

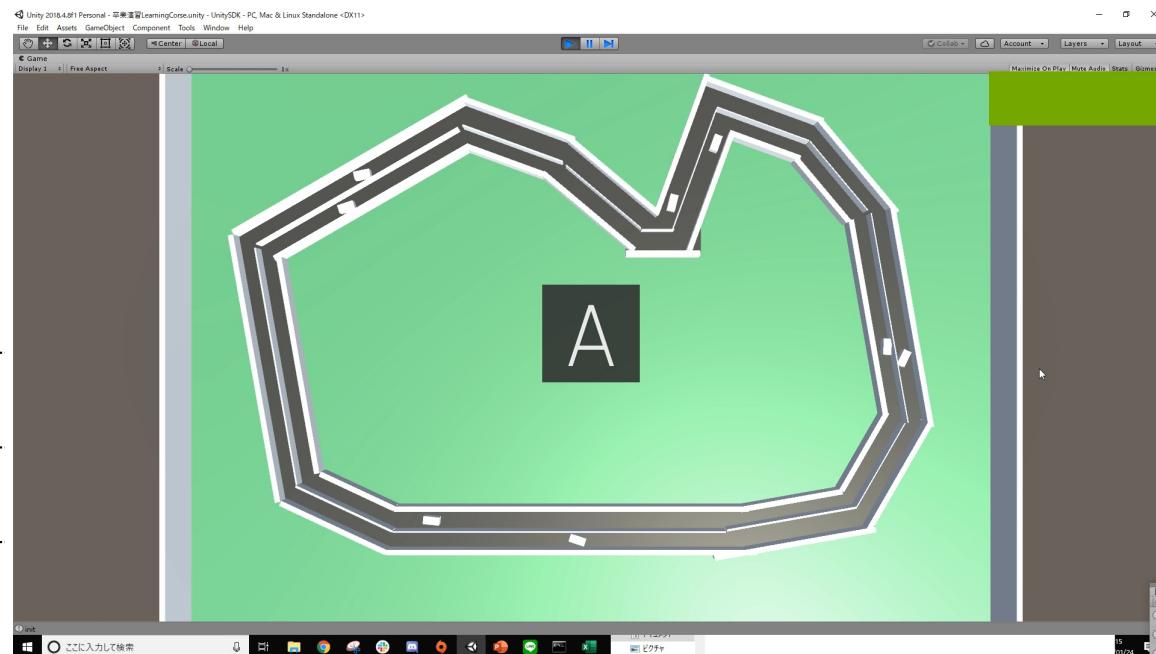
速度(0-最高速度で正規化)	+0~1
右67.5-112.5度,距離2m以内に白線が存在	+0.5
左247.5-297.5度,距離2m以内に白線が存在	-0.5
建物・車に衝突	-500

3-II-C 報酬

- 予備実験

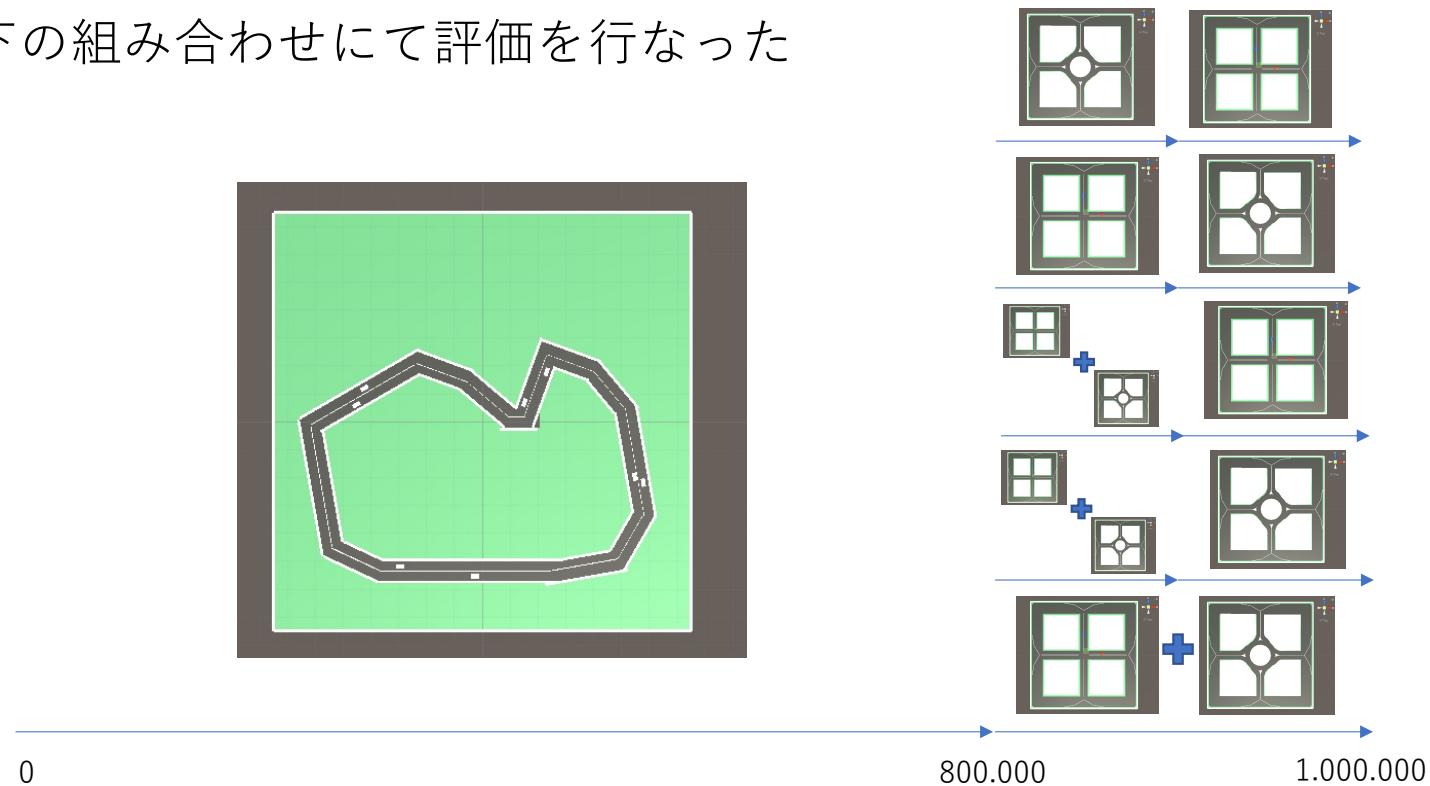
報酬によっては意図と
全く違う動きになる

速度(0-最高速度で正規化)	+0~1
右67.5-112.5度,距離2m以内に 白線が存在	+1
左247.5-297.5度,距離2m以内に 白線が存在	無し
建物・車に衝突	-500



3-III-D 学習環境組み合わせ

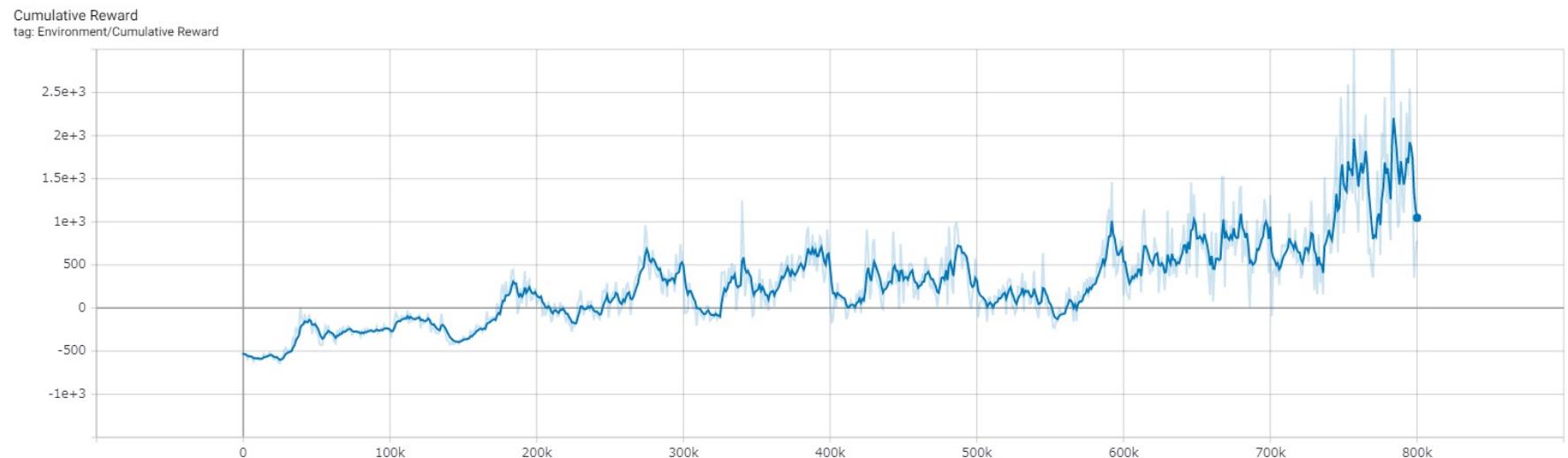
以下の組み合わせにて評価を行なった



3-II-E 学習の観測

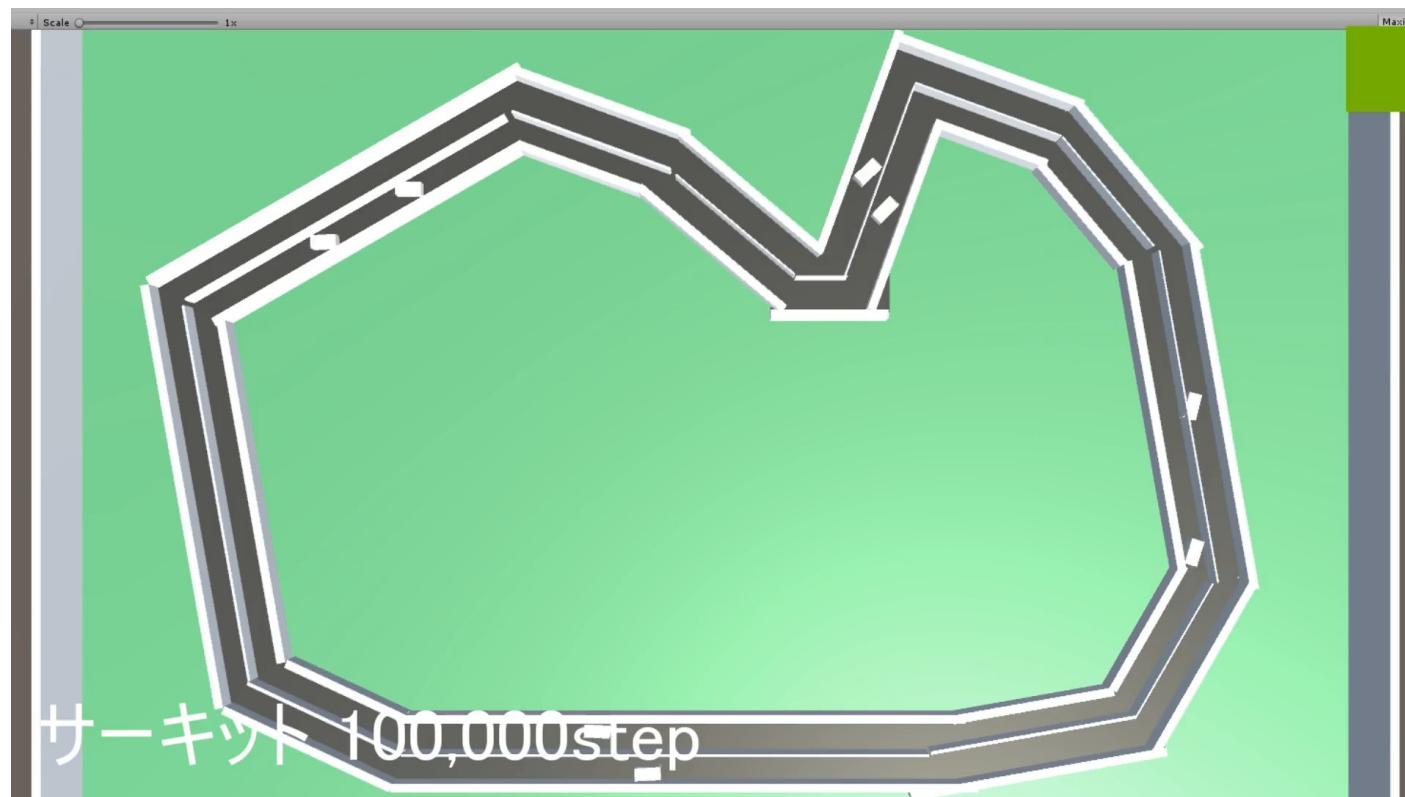
同じハイパーパラメータでも局所解に陥ってしまう事があるため、累計報酬の推移を確認しながら学習を行う

累計報酬



学習の様子(動画)

訓練用環境: サーキットでの学習経過



3-III 評価

交通システムと自動運転車を仮想環境に構築



自動運転を深層強化学習を用いて学習



生成した自動運転車を用いて仮想環境にて検証

3-III-A 評価基準

事故発生数と車両通過数 2つの指標で評価
60分間走行・測定、1分毎の平均を使用

➤ *Clash*:衝突発生数
建物・車に衝突した車数

➤ *Flow*:車両通過数
交差点を通過した車数. 交差点内で衝突した車両は含まれない

3-III-B 評価結果

Clash: 衝突発生回数

車両数	MIX		MIX→C		MIX→R		C→R		R→C	
	R	X	R	X	R	X	R	X	R	X
8	8.10	13.94	6.89	15.45	7.00	13.88	9.45	16.69	7.49	16.72
16	18.38	46.32	17.44	43.80	14.86	51.32	19.26	48.87	17.37	41.52

Flow: 車両通過数

車両数	MIX		MIX→C		MIX→R		C→R		R→C	
	R	X	R	X	R	X	R	X	R	X
8	49.10	39.70	49.10	41.65	48.73	42.11	41.30	44.30	46.96	40.72
16	57.20	37.80	53.84	36.57	62.23	34.97	43.17	44.37	51.70	42.45

R:ラウンドアバウト X:十字交差点

3-III-C 考察

Clash:衝突発生数

- 十字交差点ではラウンドアバウトに比べて **Clashが平均2.32倍**
ラウンドアバウトは深層強化学習による自動運転にとって比較的容易に走行可能と推察される
- 車両数を8から16にした際、ラウンドアバウトでは **平均2.25倍**
十字交差点では **平均3.02倍**
ラウンドアバウトは十字交差点に比べ車両密集時に走行しやすいと推察される

Flow:車両通過数

- 評価に使える指標たり得ない
交差点通過速度の指標として設定したが、評価方法として不適切と考えられる

4. 今後の課題

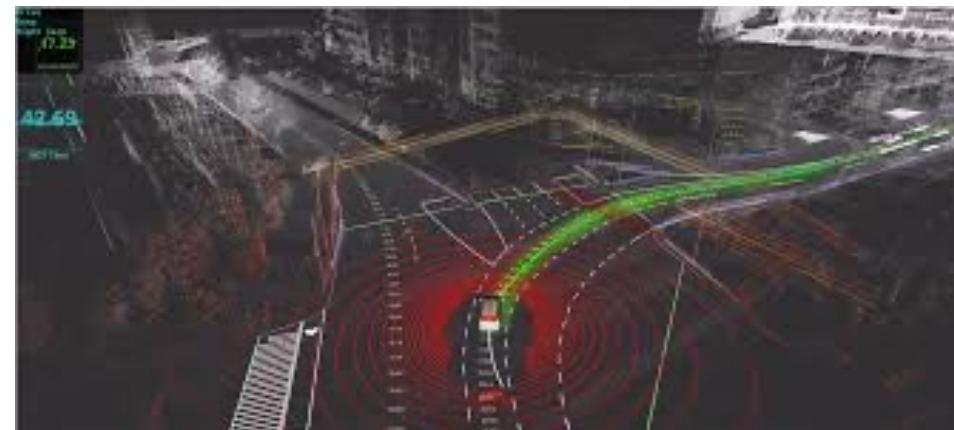
交差点内の移動速度を確認するために*Flow*を用意したが、
評価に使える指標たり得なかった
違う評価指標が必要

Flow:車両通過数

車両数	MIX		MIX→C		MIX→R		C→R		R→C	
	R	X	R	X	R	X	R	X	R	X
8	49.10	39.70	49.10	41.65	48.73	42.11	41.30	44.30	46.96	40.72
16	57.20	37.80	53.84	36.57	62.23	34.97	43.17	44.37	51.70	42.45

4. 今後の課題

- 仮想環境で構築を行ったが、あまり現実に即しているとは言はずらしい状況
自動運転OSSのAutowareを使ってより現実に近い環境で実験を行う



<https://www.pdsl.jp/fot/autoware/>