

# 自動運転車の協調型運転戦略の導入による渋滞抑制

## Cooperative Learning to Achieve Driving Strategy for Suppression of Traffic Jam

石川翔太\*<sup>1</sup> 荒井幸代\*<sup>1</sup>  
Ishikawa Shota Arai Sachiyo

\*<sup>1</sup>千葉大学大学院工学研究科都市環境システムコース  
Graduate School of Engineering, Chiba University, Division of Urban Environment Systems

An autonomous vehicle is noted to reduce traffic jam caused by perturbations in dense traffic. We focused on an automated driving vehicle's strategy where the acceleration characteristics, that is, the driving style automatically adapts to different traffic situations. Then, we propose a method to get such strategy in the multi-agent reinforcement learning. We found that phantom traffic jams are suppressed by an intelligent vehicle with the proposed strategy.

### 1. はじめに

近年, 前方の自動車の状態を観測し, 自動で速度を調整する Adaptive Cruise Control や, これに通信機能を持たせた Cooperative Adaptive Cruise Control の開発が進められている. 同時に, 速度制御が自動化されたこれらの自動車 (以下, 自動運転車と記す) が, 交通流に与える影響の解析や有効利用できる方法を考えなくてはならない.

自動運転車による渋滞緩和・抑制アプローチとして「運転戦略」が注目されている. 運転戦略には, 自動運転車が観測した状態に対する適切な速度制御が定められる. しかし, 既存研究の運転戦略 [Kesting 08, Knorr 12] はトップダウンに設計され, 最適とは限らない.

そこで, 本論文は環境との相互作用による強化学習 [Sutton 00] を用いて, 協調型の運転戦略を獲得する. 強化学習は, エージェントの試行錯誤を通して, 観測した状態に対する最適な行動を示した「方策」を獲得するアルゴリズムである. 提案手法では, エージェントを自動運転車, 方策を運転戦略として強化学習を適用する. 計算機実験により, 自然渋滞が発生する交通流において, 提案手法で獲得した運転戦略の導入が交通流に与える影響を観察する. また, エージェントの状態観測能力やエージェント間の運転戦略共有が, 学習結果および交通流に与える影響を観察する.

以下, 2章では, 関連研究をまとめ, 問題設定として自動運転車普及課程の交通流のモデル化について記す. 3章では, 強化学習の説明をした後, 運転戦略獲得への強化学習導入方法について説明する. 4章では, 計算機実験の設定, および, 実験結果を記し, 最後に5章でまとめと今後の課題を述べる.

### 2. 問題設定

#### 2.1 関連研究

高度交通システムの一環として, 道路インフラと自動車間の路車間通信, および, 自動車間の車車間通信が利用可能となってきた. これらの通信を用いた, 渋滞緩和へのアプローチをいくつか挙げる. 路車間通信を用いたアプローチは, 道路インフラにより道路の状態を観測し, 適当な運転戦略を路車間通信を使って自動車に配信する [国総研 15]. しかし, インフラの拡

張や設置するコストがかかることや設置場所を選定する課題がある. 一方, 車車間通信を用いたアプローチは, 車車間通信により局所的に得られる情報から自動車の走行ルールを変更する. Kesting らは, 交通流を5つの状況に分け, それぞれの状況に対して適当な運転戦略を定めるアプローチを提案した [Kesting 08]. 自動運転車は観測した状態から交通流の状況を推測し, 運転戦略に従って速度制御を行う. Knorr らは, 下流にいる自動車が渋滞の予兆を察知した際, その情報を車車間通信を使って上流にいる自動車に向けて発信するアプローチを示している [Knorr 12]. 情報を受信した自動車が, 車間距離を大きくして渋滞の伝播を妨げる運転戦略に従うことで, 渋滞発生を未然に防止できることを計算機実験によって明らかにしている. これらの既存研究は本論文と同様に, 円滑な交通流実現を目的として, 自動運転車が観測した状態に対する適切な速度制御を「運転戦略」として定めたアプローチである. しかし, 既存研究の運転戦略はトップダウンに設計され, 最適とは限らない. 既存研究に対して本論文では, 環境との相互作用である強化学習を用いて, 運転戦略を獲得する. 設計者はスカラー量の報酬を定義するだけでよく, 自動運転車の試行錯誤によって環境に適した運転戦略が獲得できる. つまり, 運転戦略をボトムアップに獲得する手法といえる.

#### 2.2 交通流モデル

交通流モデルに視界共有型 Nagel-Schreckenberg モデル [石川 16] を用いる. 同モデルは, 時間・空間を離散化して表現するセルオートマトンモデルのうち, 代表的な Nagel-Schreckenberg モデル [Nagel 92] の発展モデルであり, 先行車との車車間通信を用いた視界共有を導入している. 先行車の視界状態を観測できれば, 先行車の動きの予測精度が上がり, 手動運転時より小さい車間距離で速度を保てる. この動きは, レーダや車車間通信を用いて先行車の状態を観測し, 速度制御を行う自動運転車と同様の手続きを踏んでいる. 本論文では, 車車間通信による視界共有が可能な自動車を自動運転車, 不可能な自動車を手動運転車として, 自動運転車普及課程における交通流を再現する. 従来の視界共有型 Nagel-Schreckenberg モデルを一部変更し, 各自動車ごとに視界共有台数と通信可能距離のパラメータを設定可能にした. また, 車線変更ルールを導入し, 2車線の車線変更ルール [Knospe 02] を, 3車線以上でも適用可能に変更したルールを用いる.

連絡先: 石川翔太, 千葉大学大学院工学研究科, 千葉市稲毛区弥生町 1-33, 043-251-1111(代表)

### 3. 提案手法

#### 3.1 強化学習

まずはじめに、強化学習 [Sutton 00] について説明する。環境モデル  $\langle S, A, R, \pi \rangle$  と定義する。  $S$  は状態集合、  $A$  は行動集合、  $R$  は報酬関数、  $\pi$  は方策を表す。 エージェントは時刻  $t$  において、状態  $s_t \in S$  を観測し、方策  $\pi(s_t)$  に基づいて行動  $a_t \in A$  を選択する。その後、時刻  $t+1$  では  $s_t, a_t$  によって確率的に次状態  $s_{t+1}$  に遷移し、報酬関数  $R(s_{t+1})$  から生成された報酬  $r_{t+1}$  を得る。 エージェントの試行錯誤により、最終的に得られる総報酬が最大となる行動を選択する方策  $\pi$  を獲得する。

提案手法は、状態集合  $S$  と行動集合  $A$  のマッピングである行動価値関数  $Q(S, A)$  を方策に用いる Q 学習を適用する。行動価値関数  $Q$  の更新式を (1) 式に示す。ここで、学習率  $\alpha$ 、割引率  $\gamma$  はそれぞれ、学習結果を反映する割合、将来の報酬を割引して反映する割合を示す。  $a'$  は時刻  $t+1$  の状態  $s_{t+1}$  における選択可能な行動  $a_{t+1}$  のうち、行動価値関数  $Q$  が最大となる行動を示し、その時の  $Q$  値を  $\max Q(s_{t+1}, a')$  と表す。 Q 学習の方策は  $\pi(s_t)$  は  $a_t \leftarrow a'$  であり、学習が完了している時、時刻  $t$  の状態  $s_t$  において行動価値関数  $Q$  が最大となる行動を常に選択することが最適な方策となる。しかし、学習中において最適な行動を選択しただけでは局所的な行動しか選択されない。そこで、提案手法では  $\epsilon$ -greedy 選択により、探索率  $\epsilon$  でランダムに行動を選択し、  $a'$  以外の行動で探索する。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)) \quad (1)$$

#### 3.2 強化学習による運転戦略獲得

問題設定の交通流において、自動運転車に強化学習を適用し、試行錯誤の走行データを基にして運転戦略を獲得する。視界共有型 Nagel-Schreckenberg モデル [石川 16] の走行ルールは、速度変更、減速、移動の手順で構成される。ここで「減速」は、本来手動運転時に伴う速度の揺らぎを再現するための手順であり、機械的に速度制御が可能な自動運転車に適用するのは本意でない。そこで、「減速」を強化学習における「行動」とし、運転戦略を方策  $\pi$  とすることで、自動運転車に強化学習を適用する。

図 1 に、強化学習が適用された自動運転車の時刻  $t$  におけるアルゴリズムを示す。自動運転車はまず、時刻  $t$  における状態  $s_t$  を観測する。ここで、観測する状態は自分の速度・車間距離とし、視界共有を行う場合は、共有相手の速度・車間距離、共有相手との距離も観測できる。共有相手が通信可能距離の範囲外にいる時、共有相手に関する状態は全て取得不可能の 1 状態とする。次に、観測した状態  $s_t$  を基に報酬  $r_t$  を得る。報酬は速度が 0 の場合、および、車間距離が 6 を超えた場合  $r_t = -1$  の罰が与えられ、それ以外は  $r_t = 0$  とする。獲得した報酬  $r_t$  と (1) 式を基に行動価値関数  $Q$  を更新する。そして、運転戦略  $\pi(s_t)$  により、探索率  $\epsilon$  でランダムに行動を選択、それ以外は  $\max Q(s_t, a')$  となる  $a'$  で行動を選択する。

以降は、自動車の挙動に関するルールが続く。車線変更ルールは、自分の速度+1 を基準に、自分の車間距離がこの基準よりも小さく、隣車線の先行車との車間距離がこの基準よりも大きければ必ず車線変更を行う。このとき、隣車線の後続車との車間距離は考慮しないものとする。次に、走行ルールは、「1. 速度変更」、「2. 減速」、「3. 移動」の順に実行する。「1. 速度変更」では、観測した状態を基に、可能ならば加速し、加速で

```
状態観測  $s_t$ 
報酬  $r_t$  の獲得
  if (速度=0, 車間距離 >6)
     $r_t \leftarrow -1$ 
  else
     $r_t \leftarrow 0$ 
行動価値関数  $Q$  の更新
   $Q(s_{t-1}, a_{t-1})$ 
   $\leftarrow Q(s_{t-1}, a_{t-1}) + \alpha(r_t + \gamma \max_{a'} Q(s_t, a') - Q(s_{t-1}, a_{t-1}))$ 
運転戦略  $\pi(s_t)$  による行動選択
  if (1 -  $\epsilon$ )
     $a_t \leftarrow a'$ 
  else
     $a_t \leftarrow$  random select
車線変更ルールの適用
視界共有型 Nagel-Schreckenberg モデルの走行ルールの適用
  1. 速度変更
  2. 減速
    if ( $a_t ==$  減速)
      減速
  3. 移動//全ての自動車が同時に移動する
時刻  $t \rightarrow t+1$ 
```

図 1: 強化学習が適用された自動運転車の時刻  $t$  におけるアルゴリズム

きなければ最大限進行できる速度に変更する。「2. 減速」では、運転戦略  $\pi(s_t)$  で選択した行動  $a_t$  を基に、減速する/しないのどちらかを実行する。「3. 移動」では、手動運転車を含む全ての自動車が同時に移動するパラレルアップデートを行う。移動の手順終了後、時刻が  $t \rightarrow t+1$  に変化する。

手動運転車の場合、自動車の挙動に関するルールだけを実行し、走行ルールの「2. 減速」では、設定された減速確率により減速する/しないのどちらかが実行される。

#### 3.3 運転戦略共有による学習

問題設定の交通流は、学習主体となる自動運転車が複数存在する、マルチエージェント環境となる。各エージェントが独立して学習を行う場合、エージェントごとに行動価値関数  $Q$  が定義される。これに対して、運転戦略共有して学習を行う場合、全てのエージェントが一つの行動価値関数  $Q$  を同時に参照・更新する。なお、更新する順番は、交通流の上流から下流の順で行い、時刻  $t$  内で複数エージェントが同一の行動価値関数  $Q$  を更新する場合、上流にいる自動運転車の更新結果が優先される設定とした。

### 4. 計算機実験

提案法の効果を観察するための計算機実験を行い、得られた結果を示す。はじめに、運転戦略の導入が交通流量に与える影響を観察する。次に、自動運転車普及率が交通流量に与える影響を観察する。最後に、実験結果の考察を行う。

#### 4.1 実験環境

計算機実験における道路の設定を記す。本論文では、単路部ボトルネックでの渋滞を対象とする。単路部は、合流や急なカーブが無く、直線の道路部分のことを示す。この単路部に、

サグやトンネル入り口等が存在する場合、人間のドライバーによる無意識な速度の揺らぎが発生する。この減速が上流の後続車に伝播すると、やがて渋滞となる。これは自然渋滞と呼ばれ、現在の高速道路における主要原因である [国総研 15]。本論文では、道路の一部に、確率的に減速する区間を設けることでこの自然渋滞を再現する。

また、道路モデルは、一次元の周期的境界条件を適用する。この条件下では、各時刻  $t$  における道路上の車両密度が一定に保持されるため、車両間隔の均一/不均一性が交通流量に与える影響を評価しやすくなる。自然渋滞は車両密度が臨界密度を越えるか超えないかの変化によって生じるものではなく、同じ車両密度でも車両距離の過密によって生じるものである。つまり、周期的境界条件を用いることで運転戦略導入が自然渋滞に与える影響を、車両密度の影響を排除して観察することができる。

## 4.2 実験設定

全ての計算機実験は時刻  $t \rightarrow t + 1$  の変化を 1[step] とし、道路長  $L=100$ [セル]、3 車線道路、制限速度  $=5$ [セル/step] で行う。手動運転車は確率的に減速する区間 5[セル] 上を、減速確率 0.05 で減速する。自動運転車は視界共有台数 1[台]、通信可能距離 20[セル] に設定する。道路長  $L \times$  車両密度  $\rho$  の数の車両がランダムに初期配置され、その後 10000[step] 動かした後の交通流から 10000[step] 観測を行う。この観測を 1[episode] として 1 プロットにつき 10[episode] ずつ行う。実験結果の評価は、単位時間内に道路上の一地点を通過した自動車の台数を示す交通流量  $q$  を用いる。観察対象区間に存在する単位距離当たりの自動車の台数を示す車両密度  $\rho$  と、全自動車の平均速度  $\bar{V}$  を用いると、交通流量  $q$  は (2) 式によって求められる。

学習を行う場合、学習率  $\alpha=0.01$ 、割引率  $\gamma=0.9$ 、探索率  $\epsilon=0.01$  のパラメータで 200[episode] 繰り返し、その後、学習率  $\alpha=0$ 、探索率  $\epsilon=0$  とした 10[episode] の結果を用いる。

なお、観測結果は、有意水準 1[%] で検定を行い、有意な差があった時に大小の関係を認める。

$$q = \bar{V} \rho \quad (2)$$

## 4.3 実験結果

■**運転戦略の導入が交通流に与える影響** 自動運転車普及率 0.2 における、車両密度  $0.2 \leq \rho \leq 0.3$  と交通流量  $q$  の関係を観察する。図 2 に横軸を車両密度  $\rho$ 、縦軸を交通流量  $q$  とした図を示す。車両密度  $\rho \geq 0.29$  において、運転戦略共有/非共有ともに、運転戦略を導入した交通流量が、導入しない時よりも大きくなった。運転戦略共有/非共有それぞれの結果を比較すると、車両密度  $\rho \geq 0.26$  において、共有した方が、より大きな交通流量となることがわかった。

■**自動運転車普及率が交通流量に与える影響** 車両密度 0.3 における、自動運転車普及率と交通流量  $q$  の関係を観察する。図 3 と図 4 は横軸を自動運転車普及率、縦軸を交通流量  $q$  とした図を示し、自動運転車の通信可能距離はそれぞれ 20[セル] と 0[セル] に設定した。図 3 の自動運転車普及率  $\leq 0.3$  では、運転戦略を導入した方が、導入しないときよりも大きくなった。しかし、自動運転車普及率  $> 0.35$  では、運転戦略を導入しない方が交通流量が大きくなった。運転戦略共有/非共有それぞれの結果を比較すると、自動運転車普及率 0.05 を除いて、共有した方がより大きな交通流量となることがわかった。

一方、図 4 は、通信可能距離 0[セル] に設定したため、自動運転車は先行車と通信が行えず、視界共有できない状況で走行する。つまり、手動運転車と自動運転車は、確率的な減速が運

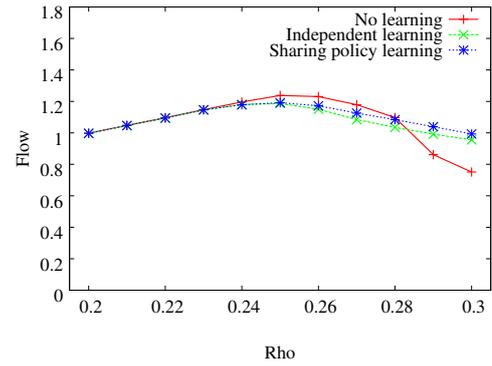


図 2: 車両密度と交通流量の関係 (自動運転車普及率 0.2, 視界共有台数 1[台], 自動運転車普及率 0.2)

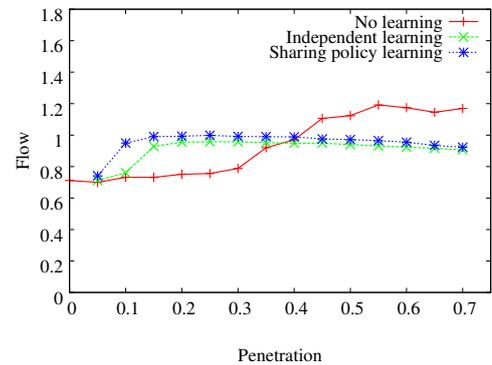


図 3: 自動運転車普及率と交通流量の関係 (車両密度 0.3, 視界共有台数 1[台], 通信可能距離 20[セル])

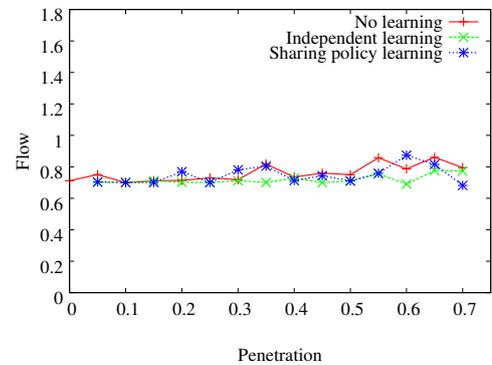


図 4: 自動運転車普及率と交通流量の関係 (車両密度 0.3, 視界共有台数 1[台], 通信可能距離 0[セル])

転戦略による減速かの違いだけとなる。このとき、図 4 の通り、運転戦略導入/非導入、運転戦略共有/非共有が交通流量に与える影響はわからなかった。

## 4.4 考察

■**運転戦略の導入が有効となる状況** 図 2 と図 3 の結果より、車両密度  $\rho \geq 0.29$  の時、および、自動運転車普及率  $\leq 0.3$  の時に提案手法の運転戦略導入が有効であることがわかった。この理由を、停止ステップ数から考察する。図 5 では、横軸を車両密度  $\rho$ 、縦軸を 1[episode] における自動車 1[台] 当たりの平均停止ステップ数 [step] とした図を示す。図より、運転戦略を導入していない交通流において、車両密度  $\rho=0.27$  から平均停

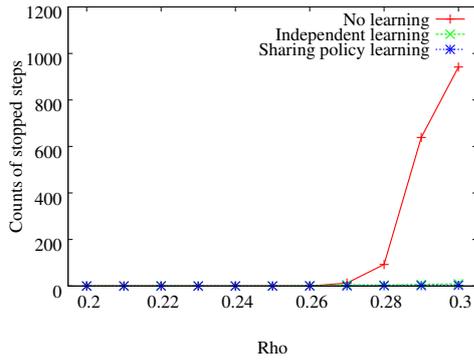


図 5: 車両密度と 1[episode] における自動車 1[台] 当たりの平均停止ステップ数 [step] の関係 (自動運転車普及率 0.2, 視界共有台数 1[台], 通信可能距離 20[セル])

止ステップ数が増加しはじめ、 $\rho=0.3$  では約 941[step] となることが確認できる。これに対して車両密度  $\rho \leq 0.26$  では、平均停止ステップ数が 1[step] 未満である。自動運転車は停止した場合に罰が与えられるため、車両密度  $\rho \leq 0.28$  では罰が与えられる機会がほとんど無いと考えられる。従って、運転戦略獲得までの十分な学習機会が得られず、学習が完了していない運転戦略を導入したことにより交通流量が小さくなったと考えられる。自動車 1[台] 当たりの平均停止ステップ数が大きくなる状況、つまり、自然渋滞が発生している状況では、設定した罰で学習する運転戦略が有効となる。

■視界共有と運転戦略共有が学習に与える影響の考察 図 3 と図 4 を比較すると、自動運転車普及率 0.05 以外は、視界共有した方が、視界共有しないときよりも大きくなった。これは、先行車の視界情報が加わることで観測可能な状態数が増加し、より多くの状態に対する行動が示された運転戦略を獲得できたためである。

図 3 の結果より、自動運転車普及率 0.05 以外は、運転戦略共有する学習の方が、独立した学習よりも交通流量が大きくなった。この結果を考察するために、図 6 に、横軸をエピソード数、縦軸を交通流量  $q$  とした学習曲線の図を示す。エピソード初期 (1[episode] から 20[episode] まで) において、独立して学習するよりも、運転戦略共有して学習した方が早く交通流量が大きくなっている。これは、運転戦略共有により複数エージェントで行動価値関数  $Q$  を更新したためだと考察する。また、独立して学習する場合の学習曲線は収束していないことがわかる。そのため、運転戦略共有/非共有それぞれの学習の、どちらがより有効な運転戦略が獲得できるのかは、200[episode] 以上のエピソード数で学習が収束するのを確認した結果を用いて判断する必要がある。

## 5. まとめと今後の課題

本論文は、運転戦略による交通流の最適化を目的とし、強化学習を用いて獲得した運転戦略が交通流に与える影響を観察した。トップダウンに設計されていた既存の運転戦略に対して、提案手法は自動運転車の試行錯誤によってボトムアップに運転戦略を獲得することができる。自動運転車を学習主体として、自動運転車が停止した場合および一定以上の車間距離となった場合に罰を与えることで、自動運転車は自らが停止しないかつ一定以下の車間距離を保てる運転戦略を学習する。計算機実験の結果、自然渋滞が発生する交通流において、提案法の運転戦略を導入すれば、導入しないときよりも交通流量が大きくなる

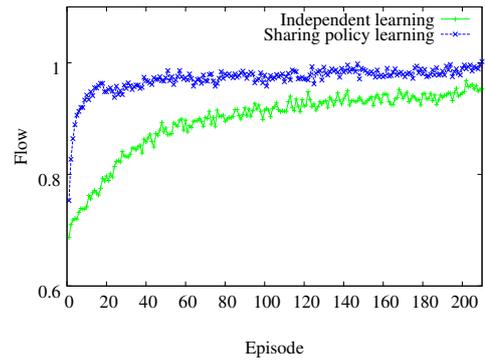


図 6: 独立した学習と運転戦略共有した学習時の学習曲線 (自動運転車普及率 0.2, 視界共有台数 1[台], 通信可能距離 20[セル])

ことが示された。また、自動運転車は視界および運転戦略共有をすることにより、運転戦略導入時の交通流量が大きくなることや、学習の収束が早くなることがわかった。

今後の課題に、設計した報酬と交通流量の関係を観察することを挙げる。

## 参考文献

- [石川 16] 石川翔太, 荒井幸代: 先行車情報の共有が自然渋滞に与える影響の解析-Nagel-Schreckenberg Model の一般化-, 人工知能学会論文誌, Vol.31, No.2(2016 掲載予定)
- [Kesting 08] Kesting, A., Treiber, M., Schönhof, M., and Helbing, D.: Adaptive cruise control design for active congestion avoidance, *Transportation Research Part C, Emerging Technologies*, Vol.16, No.6, pp.668–683(2008)
- [Knorr 12] Knorr, F., Baselt, D., Schreckenberg, M., and Mauve, M.: Reducing traffic jams via VANETs, *Vehicular Technology, IEEE Transactions on*, Vol.61, No.8, pp.3490–3498(2012)
- [Knospe 02] Knospe, W., Santen, L., Schadschneider, A., and Schreckenberg, M.: A realistic two-lane traffic model for highway traffic, *Journal of Physics A: Mathematical and General*, Vol.35, No.15, pp.3369–3388(2002)
- [国総研 15] 高速道路サグ部等交通円滑化研究会: 「高速道路サグ部等の渋滞箇所への効果的な交通円滑化対策の実現に向けて」中間とりまとめ, 国土交通省, 国土技術政策総合研究所 (2015)
- [Nagel 92] Nagel, K., and Schreckenberg, M.: A cellular automaton model for freeway traffic, *J. Phys. I France*, Vol.2, No.12, pp.2221–2229(1992)
- [Sutton 00] Sutton, R. S. and Barto, A. G.: 強化学習, 森北出版 (2000)