

強化学習を利用した信号制御 Traffic Signal Control by Reinforcement Learning

後藤 健志[†] 鈴木 輝彦[‡] 太原 育夫[‡]
Kenji Goto Teruhiko Suzuki Ikuo Tahara

1. はじめに

近年、個人の所有する車両数が増加し、これに従って交通流の複雑化、増加が招かれている。交通渋滞の発生による経済的損失や二酸化炭素の増加原因の一因にもなっている排気ガスの増加といった問題が発生するため、交通渋滞の解消が必要とされている。その手段の一つとして信号制御システムの効率化が挙げられる。現在、信号制御方式は単独や中央管理方式、系統制御が主に用いられている。しかし、これらの方式は扱う交差点数が増えるほど、計算コストの増加による制御機能の低下が現れて、更にはそれらの信号機の改良や維持には多くの費用を要することになる。本稿では、道路上の交差点に信号エージェントを配置して、交通権の切り替えに強化学習を用いて制御することで、交通信号の効率化を試み、その評価を行う。

2. Q 学習を用いた信号制御

本稿では信号機をエージェントとみなした信号制御システムを提案する。信号制御には Q 学習を用いる。

信号制御の手順は以下の通りを行う。

- (1) 各現示の開始時、信号エージェントは環境を認識する
- (2) 認識した環境から信号エージェントはあらかじめ決めておいた時間の中から ϵ -greedy 手法に従って、青時間を選択する
- (3) 選択した時間経過後、評価関数の評価により選択した青時間の評価を報酬で与え、その時間の Q 値を更新していく
- (4) (1)~(3) を繰り返して行う

個々のエージェントが環境として知覚できる情報は以下の2種類とする。

- ・自分の交差点流入路上に存在する車の数
- ・自分の交差点を通過する車両台数

各エージェントの状態と行動は以下の通りである。

- 状態集合 $S = H \times \text{sidestate}$
H 流入路上に存在する車数をその台数に応じて、0,1,2,3 に対応させる。状態数 4^4
sidestate 現在、どの方向に青現示を与えているか
状態数 2
- 行動集合 A あらかじめ定めた七つの青時間
8,16,24,32,40,48,56[step]

さて、信号制御を行うに当たり学習の対象を何とするかが問題になる。制御パラメーターを直接学習する事が

第一に考えられるが、パラメータと結果の関係解析が困難である。そこで、それぞれの現示のときにどれだけの青時間を取ればよいのかを学習するものとした。

報酬 R は以下のように流す方向の車の台数に対して適切な青時間を選択したかによって与えられる。現示終了時に信号エージェントは観測した結果から評価関数 $Observe(t)$ の値を決定する。

$Observe(t) = \text{Bluetime}(t) + \text{throwcarnum}(t) \times \text{throwtime}$

ここで $\text{throwcarnum}(t)$ 流入台数

$\text{Bluetime}(t)$ 選択した青時間

throwtime 車1台が通過後に次の車が来るまでの間隔

であり、 t は系内での時間を表す。

このとき、

$$R = \begin{cases} 1 & (\alpha \leq Observe(t) \leq \beta) \\ -1 & (else) \end{cases}$$

とする。

ここで

$$\alpha : \text{Waitcar}(t) \times \text{throwtime} \times a$$

$$\beta : \text{Waitcar}(t) \times \text{throwtime} \times b$$

となる。

ただし、 $\text{Waitcar}(t)$ は各現示の通行可能方向で流入路上に存在する車が多い流入路上の車の台数であり、 a, b は定数とする。

3. 実験

提案手法の有効性を検証するためシミュレーション実験を行った。シミュレーションを行う環境は、道路形状を碁盤目状、交差点数を 5×5 、交差点間隔が 30 セル、車線を片側 1 車線 (右折レーンあり) とした。また、道路上の車の動作を ASEP モデル [3][4] を利用してシミュレーター上に再現した。車の動作は車の長さを 5m、前進確率を 0.9 とした。また、評価関数を挟むための α, β 中の a, b を $a = 1.5, b = 3.0$ とした。そして、Q 学習で用いる学習率を 0.1、割引率を 0.9 と設定した。

3.1 実験概要

この実験では、現実的に起きる交通状態を模して以下の2つの交通状態の元で実験を行った [5]。

また、車は外部に接続する道路の端から一定間隔で発生して、外部に接続する道路の端に車が到達した時点で車は消滅するものとする。

実験 1 では 10000step、実験 2 は 20000step まで時間を設定して実験を行った。

- 実験 1 東西と南北の流入量に差がある状態
流入率 $4/5$ [台/step] で東方向からの流入確率を 75%、

[†]東京理科大学大学院理工学研究科情報科学専攻

[‡]東京理科大学理工学部情報科学科

南方向からの流入確率を25%と設定し、入ってきた道路から直進のみを行う。

- 実験2 交通流の変化
流入率を

$$\begin{cases} 4/4[\text{台}/\text{step}] & (4000\sim 8000, 12000\sim 16000) \\ 4/7[\text{台}/\text{step}] & (\text{else}) \end{cases}$$

で変化させた場合をみる。

今回の信号機としては、サイクル長80step スプリット0.5で固定したNormal-signalと単独で強化学習を行い、信号制御していく信号機Proposal-signalを用いて比較を行った。

この実験においては右折車に関する制御を考慮に入れていないので、極端に車の台数が増えた時に右折車の処理が出来なくなり、正当な評価が出来なくなる可能性がある。そこで、右折車が発生する確率を極端に少なく設定してある。

3.2 実験結果

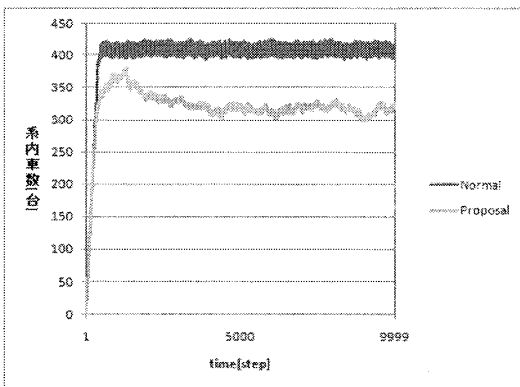


図1: 実験1での系内車数

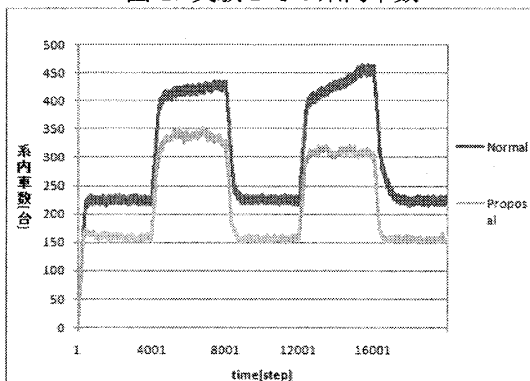


図2: 実験2での系内車数

実験1での系内車数を図1に、実験2での系内車数を図2に示す。実験図の縦軸は系内車数[台]、横軸はtime[step]である。ここで系内車数とはシミュレーター上の道路に存在する全車台数である。

3.3 実験結果と評価

実験1について、図1よりProposal-signalは始めの1500stepまでは系内車数が増加傾向にあったが学習が進むにつれて一定の値に収束するのが確認できる。Proposal-signalは交通量が少ない現示には少ない青時間を割り振るので結果として交通量が多い現示に多くの青時間が割り振られるようになりスムーズに交通が行われたと言える。

実験2については、図2より系内車数に注目すると8000stepでの交通量減少の際はNormal-signalとProposal-signalが同程度の時間で再び元の安定した車数に戻っているが、16000stepでの再度起きた交通量減少の際にはProposal-signalの方がNormal-signalよりも早く元の車数に戻っており、学習の効果が現れていると言える。

Normal-signalは交通量の変化に対応できず、8000stepより16000stepでの交通量の増加の方が系内車数増加の傾向がみられる。これは1回目の増加時に処理しきれなかった車数の影響が残り続け、2回目の増加時にはその処理しきれなかった車の台数だけ系内車数の増加につながっていると考えられる。

4. おわりに

本稿では、個々の信号機の制御に強化学習を用いて制御を行う方式を提案した。単独の信号だけでも交通流の変化に合わせて適切な青時間を選択する事が可能であり、様々な交通状況を試してみた結果、固定方式よりも良い結果を得られた。

提案手法では、交通流の変化にも対応する事が可能であり、改めてパラメーターをやる必要がなくなり、信号の改良や維持への負担の軽減に貢献できたと言える。

また、ここでは紹介していないが、渋滞が発生する交通流があり、単独の信号機だけでは処理しきれない場合、複数の信号群が互いに連携する仕組みが必要であると考えられる。

参考文献

- [1] 社団法人 交通工学研究会, 改訂 交通信号の手引き, 丸善株式会社, 2006.
- [2] 高玉樹, マルチエージェント学習-相互作用の謎に迫る-, コロナ社, 2003.
- [3] 西成活裕, 渋滞学, 新潮社, 2006.
- [4] 内田英明, 荒井幸代, “複車線交通流に与える運転選択速度と後方視野の影響,” 第23回人工知能学会全国大会論文集, 2I2-2, 2009.
- [5] 参沢匡将, 木村春彦, 広瀬貞樹, 大里延康, “強化学習型マルチエージェントによる交通信号制御,” 電子情報通信学会論文誌 D-I, Vol. J83-D-I, No.5, pp.478-486, 2000.