

一般道における渋滞緩和・解消を図る 交通信号機制御システムに関する研究

榎本大義[†] 北越 大輔[†] 鈴木 雅人[†]

東京工業高等専門学校[†]

1. はじめに

近年、特に都市部で頻発する交通渋滞は、時間的損失をはじめとする様々な問題の要因として指摘され、社会問題となっている。交通渋滞を解消する一アプローチとして、一般道における各交差点をエージェントとみなした強化学習法によって、信号点灯間隔の最適化を目指す研究がある[1]。本研究では文献[1]で提案された渋滞緩和・解消システムにおけるエージェント間の学習格差を抑制するため、宮内らにより提案された相互作用型階層強化学習システム (IH-RL) [2]の概念を当該システムへ導入し、渋滞緩和性能の改善を図る。仮想道路環境に対して提案システムを適用し、性能を検証する。

2. 仮想道路環境

本稿で用いる仮想道路環境 (図 1) は、[1]における設定を参考としている。仮想環境は、交差点・終端点 (車の流入, 消失点)・道路で構成される。各交差点に設置する信号には、青信号と右折車用矢印を東西・南北各 2 方向に設け、計 4 種類の点灯状態を設定する。当該環境では次の (1)~(3)に示す処理を 1 ステップとする：
 (1) 各終端点に設定された流入率 (車の増加台数/ステップ) に応じ、ポアソン分布に従った確率で隣接道路に車を生成 (2) 現在の信号の点灯状態に合わせ、各交差点の通過待ち行列にいる車を次の道路へ移動 (3) 各道路上の車を、車の速度に比例した確率で進行方向の交差点待ち行列へ移動。1 ステップは実世界における 2 秒を想定している。信号の各点灯状態を 1 周するため要するステップ数を 1 単位時間とする。[1]で設定された仮想環境では、一般道における渋滞を模倣し

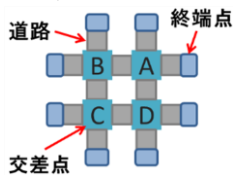


図 1 仮想道路環境

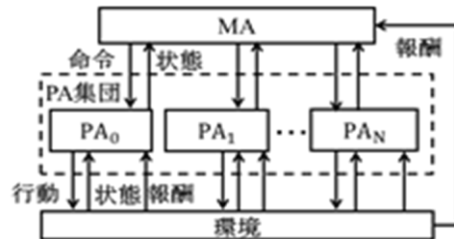


図 2 IH-RL の構成

た一方、その定義は十分でなく、渋滞の緩和度合いに関する評価が困難であったため、本稿では渋滞の定量的定義を定め、提案システムの特徴・性能の具体的な評価のために活用する。

【本稿における渋滞の定義】 仮想環境上の全道路に、存在できる上限車両数 N を設定し、 $N/2$ を通過待ち行列に到達できる上限車両数とする。ここで、「通過待ち行列に存在する車の数 $\geq N/2$ 」となった状態を渋滞と定義する。

3. 相互作用型階層強化学習システム (IH-RL)

強化学習エージェントは、報酬を手掛かりに試行錯誤的な行動を通して環境に適応した振舞 (方策) を学習する。エージェントは観測状態 s および s において出力する行動 a の対に対して与えられた重み $w(s, a)$ を用いて行動選択し、重みを更新することで方策学習を進める。

IH-RL は、マルチエージェント環境下でエージェント群が協調的行動を獲得可能な学習システムである。IH-RL では Playing Agent (PA) と Management Agent (MA) が存在し、各々が強化学習で方策を学習する。PA は自身の目的を達成すべく方策を学習し、MA は、PA 集団の学習状況を入力として、集団全体の目標を達成する方策 (PA 集団に対する適切な命令の与え方) を学習する (図 2)。PA と MA にはそれぞれ Profit Sharing (PS), Bucket Brigade (BB) と呼ばれる強化学習法を適用する。時刻 $t = L$ において目標状態に達し報酬 r を獲得した際、PS におけるルール重みは(1)式によって更新される ($\gamma \in ([0, 1])$: 減衰率)。

$$w(s_t, a_t) = w(s_t, a_t) + r \cdot \gamma^{L-t} \quad (t = 0, 1, \dots, L) \quad (1)$$

一方、BB では t 回目の行動選択後に報酬獲得時には(2)式、未獲得時は(3)式に従いルール重みを

A study on Traffic Signal Control System to Alleviate Traffic Congestion in General Road Environment

[†]Hiro Yoshi Enomoto, Daisuke Kitakoshi, Masato Suzuki, Tokyo National College of Technology

更新する ($\beta \in ([0,1])$: 学習速度に関する定数)。

$$w(s_t, a_t) = (1 - \beta) \cdot w(s_t, a_t) + r \quad (2)$$

$$w(s_t, a_t) = (1 - \beta) \cdot w(s_t, a_t) + \beta \cdot w(s_{t+1}, a_{t+1}) \quad (3)$$

4. IH-RL を用いた渋滞緩和・解消システム

本節では、IH-RL を用いた渋滞緩和・解消システムの基本的枠組みについて紹介する。

4-1. PA に関する設定

仮想環境中の各交差点を IH-RL における各 PA とする。PA の状態は、自身と隣接する東西・南北 2 方向の道路における通過待ち車数 w_H , w_V と信号能力 (一度の青信号点灯で通過させられる車の数) a_H , a_V との比 s_H , s_V ($s_H = w_H / a_H$, $s_V = w_V / a_V$) と、その変化量 c_H , c_V から決まる。 s_H , s_V , c_H , c_V は閾値によってそれぞれ 3 値 ($\{0,1,2\}$) に量子化されるため、PA の総状態数は 81 となる。 s_H 及び s_V が 1.0 より大きい時は、信号能力が不足していることを意味し、これらの値が 1.0 と等しい時は、交通需要と信号能力が調和している状態を意味する。ただし、周囲の道路状況によっては、必ずしも $s_H = s_V = 1.0$ が適切であるとは限らないため、 s_H と s_V に対して目標値 S_A (初期値 $S_A: 1.0$) を用意する。また、PA は自身の有する 4 種すべての信号における点灯時間の増減 $\{+3,0,-3\}$ (単位: ステップ) を行動とする。PA は単位時間毎に状態観測および行動選択を繰返し、状態観測時において $|s_H - S_A| < \epsilon$, および $|s_V - S_A| < \epsilon$ の時に報酬 $r = 100$ を獲得する (ϵ : 定数)。

4-2. MA に関する設定

MA は 400 ステップ毎に状態観測と行動選択を行う。MA の状態は、MA の $n - 1 \sim n$ 回目の行動選択までにおける PA_i ($i = 0,1, \dots, 3$) の通過待ち車数の平均値 M_i から決まり、3 値 ($\{0,1,2\}$) に量子化されるため、総状態数は 81 となる。また、一体の PA_i に対する S_{A_i} の増減 $\{+0.05, 0, -0.05\}$ を行動とする。MA は n 回目の行動選択時における M_i の標準偏差 σ_n について、 $\sigma_{n-1} > \sigma_n$ の時に報酬 $r = 100$ を獲得する。

5. 計算機実験および結果・考察

提案システムの比較対象として、MA 導入前のシステム (文献[1]におけるシステム. 以降従来システムと呼ぶ) を準備し各 10 回の実験を実施する。実験は、1 回あたり 8,000,000 ステップで終了する。1 交差点における平均待機車数の推移を図 3 に示す。図から、双方のシステムにおける待機車数の減少が確認でき、ステップ数の

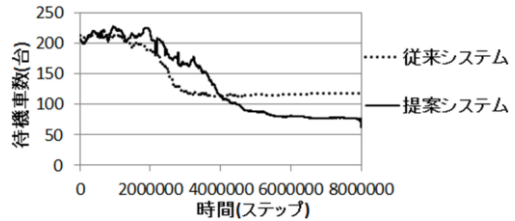


図 3 1 交差点における平均待機車数

表 1 実験後半部における σ_n の平均

	σ_n の平均
従来システム	93.11
提案システム	47.18

増加に伴い提案システムにおける値がより小さくなっていることから、当該システムがより効果的に渋滞を緩和しているといえる。これは、MA の学習によって PA 間の格差が抑制されているためと考えられ、表 1 に示す結果からも確認できる。また、実験終盤 (6,000,000 ステップ以降) において提案システムでは渋滞が発生しないことも確認した。従来システムにおける渋滞発生直前の道路状況を調査したところ、渋滞道路に隣接する道路の待機車数は渋滞道路より多く、時間が経過すれば渋滞道路へ多量の車が流入する状況であった。渋滞道路を制御する PA は、渋滞の発生に備えて当該信号の能力を強化する必要があるが、提案システムでは、MA が PA 全体の状況を把握しながら適切に S_A を変動させることで、各 PA の信号能力を調整し、突発的な渋滞を回避出来たと考えられる。

6. まとめ

本稿では仮想一般道環境において、IH-RL を用いた渋滞緩和・解消システムを提案し、計算機実験によってその性能を評価した。実験の結果、IH-RL の導入によって交差点における待機車数の減少、突発的な渋滞発生回避が可能となった。今後の課題としては、より効果的な IH-RL の学習設定の検討、仮想環境の拡張などが挙げられる。

参考文献

- [1] 山崎大地, 北越大輔, 鈴木雅人 “仮想道路環境上における強化学習による交通渋滞緩和・解消システムの性能評価”, 知能システムシンポジウム予稿集, pp. 23-26, 2010.
- [2] 宮内龍之介, 北越大輔, 鈴木雅人, “相互作用型階層強化学習システムを用いたエージェント群の協調行動獲得”, FAN2009 予稿集, pp. 484-489, 2009.