

自動運転車間の方策共有による分散協調型の交通流最適化

石川 翔太† 荒井 幸代†

千葉大学大学院 工学研究科 建築・都市科学専攻†

1. はじめに

市場に普及し始めた自動運転車の主な目的は、運転補助や安全確保であり、渋滞緩和・抑制に向けた交通流最適化ではない。しかし、自動運転による操作は「どんな状況でどんな走行(加減速)をするか」ルールさえ明らかになれば人よりも精緻な速度制御が可能である。そこで、本研究では交通流の最適化の実現を目的とした運転戦略の獲得を考える。

既存研究の運転戦略[1]はヒューリスティクスであるため最適な戦略とは限らない。これに対して提案手法では、自動運転車の走行情報を基にした学習により運転戦略を獲得する。具体的には、各自動運転車が運転戦略(以後方策と記す)を共有しながら、走行情報を基に方策を更新し、更新結果を再び共有することを繰り返す。計算機実験の結果、提案手法の方策を導入することで、導入しなかった時よりも交通流量、停止時間が改善した。

2. 問題設定

自動運転車普及過程における自動運転車と手動運転車が混在した交通流はセルオートマトンモデルの一つである視界共有型 Nagel-Schreckenberg[2]を基に再現する。各自動車に対して、視界共有できる自動車とできない自動車を設定可能にして、それぞれを自動運転車と手動運転車とした。また、文献[3]を基に複数車線における車線変更ルールを導入した。自動運転車の場合、前を走る自動運転車と視界共有を行いながら走行しつつ、与えられた方策に従って加減速を行う。手動運転車の場合、道路上の一部区間で確率的に減速し、手動運転時に発生する「速度の揺らぎ」を再現している。

図1に、横軸を自動運転車普及率、縦軸の左側を交通流量、右側を一台当たりの停止時間とした、方策を導入しない時の視界共有型 Nagel-Schreckenberg のグラフを示す。普及率が大きくなるにつれて、交通流量は大きく、停止時間は小さくなっている。普及率0から0.35までの間、交通流量に変化が無いことから、自動運転車普

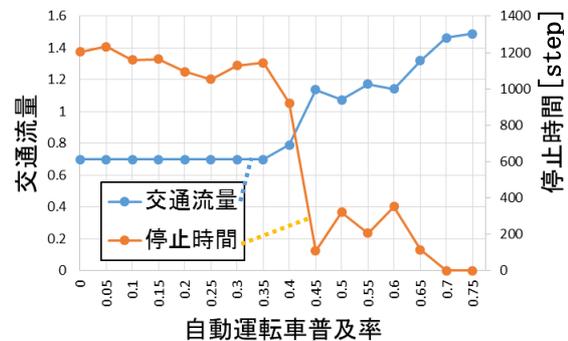


図1 視界共有型 Nagel-Schreckenberg による自動運転車普及率と交通流量・停止時間の関係



図2 方策の導入方法の概観

及過程の、特に初期段階における方策の導入が交通流を改善すると考えられる。

3. 提案手法

図2に方策の導入方法の概観を示す、自動運転車は一定の間隔に設置された路車間通信インフラから、その間隔内における方策を受けとる。本論文ではこの方策を学習によって獲得する。自動車の行動が後から遅れて交通流に影響を及ぼす点、および、手動運転車に不確実性が存在する点を考慮して、強化学習(Q学習)[4]を用いる。

3-1. 強化学習

自動運転車を学習主体のエージェント、環境モデルを (S, A, R, π) と定義し、 S は状態集合、 A は行動集合、 R は報酬関数、 π 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 π に基づいて行動 $a_t \in A$ を選択する。その後、時刻 $t+1$ では次状態 s_{t+1} に遷移し、報酬 $r_t \in R$ を得る。この繰り返しの中でQ学習では行動価値関数 $Q(s_t, a_t)$ を式(1)によって更新する。行動は ϵ -greedy 選択により探索率 ϵ でランダム行動、それ以外は $\pi(\max_a Q(s_t, a))$ として状態 s_t で Q が最大となる行動をとる。ここで α は学習率、 γ は割引率で、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。

Traffic Optimization via Decentralized Control among Autonomous and Cooperative Vehicles

Shota Ishikawa†, Sachiyo ARAI†

†Department of Urban Environmental Systems Course, Graduate School of Engineering, Chiba University

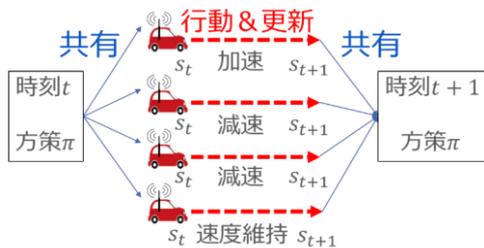


図 3 方策共有による行動価値関数の更新の流れ

表 1 計算機実験の結果

	方策無し	方策あり
交通流量	0.7	1.0
停止時間[step]	1094.1	2.7

$$Q(s_t, a_t) \leftarrow$$

$$Q(s_t, a_t) + \alpha \left[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right] \quad (1)$$

3-2. 方策共有

本問題は、複数の学習主体が存在するマルチエージェント環境であるため、エージェント間の分散協調による交通流制御が必要である。分散協調を実現するために、方策 π をエージェント間で共有しながら学習を行う。図3に示す通り、各エージェントは共有した方策に従って行動を決定し、式(1)を更新する。また更新結果を時刻 $t+1$ で再び共有する。すなわち、あるエージェントの更新結果を全てのエージェントが参照し利用可能となる。

4. 計算機実験

4-1. エージェントの状態・行動・報酬

エージェントが時刻 t で観測できる状態 s_t は「自分の速度・車間距離、通信相手との距離、通信相手の速度・車間距離」の全49392状態とし、そのときの行動は「減速する・しない」の2行動である。報酬 r_t は自動車停止した時、または、車線変更によって前に割り込まれるほどの車間距離を空けた時に負の報酬を与える。

4-2. 実験結果および考察

道路長100[セル](確率的減速が適用されるセル5[セル])の3[車線]、減速確率0.05、車両密度0.3、自動運転車普及率0.2に設定した。周期的境界条件で交通状況(車両密度と自動運転車普及率)を固定して効率よく同状況の学習を可能にした。自動車は制限速度5[セル/step]、自動運転車は視界共有台数2[台]、車車間通信可能範囲100[セル]とした。10000[step]を1[エピソード]として200[エピソード]繰り返す。学習率 α は0.01、割引率 γ は0.9、探索率 ϵ は0.01に設定した。最後の10[エピソード]は学習率 α と探索率 ϵ

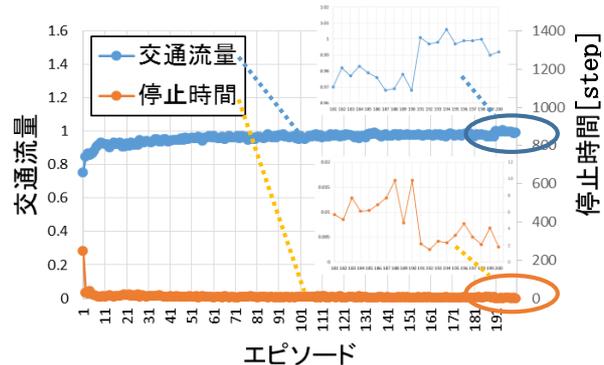


図 4 学習曲線(最後の20[エピソード]時の交通流量の拡大図を右上に、停止時間を右下に示す)

を0に設定し、このときの交通流量および一台当たりの停止時間の平均値を実験結果として表1に示す。

表1が示す通り提案手法の方策導入によって、交通流量、停止時間ともに改善した。特に停止時間は0に近く、渋滞の抑制ができたと考えられる。

図4に、横軸をエピソード、縦軸の左側を交通流量、右側を1台当たりの停止時間とした、学習曲線を示す。交通流量、停止時間ともにエピソードの初期段階から徐々に改善され、最終的に収束している。

5. まとめと今後の課題

本研究では、自動運転車の運転戦略導入による交通流最適化を目的として、強化学習を用いた運転戦略の獲得法を提案した。提案手法ではマルチエージェント環境下で、全エージェントが方策共有した状況で学習した。計算機実験の結果、提案手法によって獲得した運転戦略を導入した場合、しない場合と比較して交通流量と停止時間ともに改善した。

今後の課題として、合流や車線減少などの様々な交通状況に対する提案手法の適用可能性を調べる。また、マルチエージェント環境下で、方策共有する強化学習の解の収束性や最適性を理論的側面からの検討することを挙げる。

参考文献

[1] 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)
 [2] 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)
 [3] Rickert, M., Nagel, K., Schreckenberg, M., and Latour, A.: Two lane traffic simulations using cellular automata, Physica A, Vol.231, No.4, pp.534-550 (1996)
 [4] Sutton, R. S. and Barto, A. G.: 強化学習, 森北出版 (2000)