The Fourteenth International Symposium on Artificial Life and Robotics 2009 (AROB 14th '09),
B-Con Plaza, Beppu, Oita, Japan, February 5 - 7, 2009

Modeling of Traffic Flow using Cellular Automata and Traffic Signal Control by Q-learning

Kiyoshi Umemoto, Ji-Sun Shin, Tomofumi Ohshita, Yohei Osuki (Waseda Univ.),
Michio Miyazaki (Kanto Gakuin Univ.), and Hee-Hyol Lee (Waseda Univ.)

(Tel: 093-692-5017; Fax: 093-692-5021)
(Email address: tumsk_osaka@yahoo.co.jp)

Abstract: Recently, the flow of traffic has increased in the cities, and it has caused problems because of CO_2 emissions due to traffic jams. The traffic signal control is a typical counter measures for the congestion easing. The traffic signal control method includes the point control, the series control, and the wide area control, and the cycle time, the split, and the offset are used as the control parameters of the traffic signal. The offset is the difference of the start for the green signal between adjoining crossroads. The existing researches to generate the offset automatically are the cycle-less control technique, the real-time simulation using GA, and the optimization technique by the inclination method.

First, the traffic flow is modeled to reproduce the movement of the vehicle on the road in this paper. There are two models of the traffic flow being developed now: one is to model the traffic flow as a continuous style, and the other is to regard the vehicle as the individual movement and to form the whole flow. The traffic flow is modeled using the cellular automata as the latter case here. A traffic signal is consisted as an agent and the agent learns the control parameters of the traffic signal, which are the split and the offset under the fixed cycle length, using Q-learning method. In this paper, the offset of the signal agent is deduced using Q-learning method considering the adaptation for the dynamic change of the traffic flow.

Keywords: Traffic Flow, Traffic Signal Control, Offset, Q-learning, Cellular Automata, Agent

I. INTRODUCTION

The traffic signal control is a typical counter measures for the congestion easing. The traffic signal control method includes the point control, the series control, and the wide area control, and the cycle time, the split, and the offset are used as the control parameters of the traffic signal. The traffic signals are generated by an autonomous agent, and each signal agent studies the control parameter independently.

In this paper, a traffic signal control by the multi agent type is proposed. The traffic flow of the road network including crossroads is modeled by the traffic flow cellular automata. Under the cycle of the signal interval is constant, the control parameter of the signal agent are adjusted by Q-learning method.

II. MODELING OF TRAFFIC FLOW USING CELLULAR AUTOMATA

As the model of the traffic flow, a typical Elementary CA [1] of Wolfram by one dimension cellular automata is developed to a two dimensional model including crossroads here. The road is divided into two or more cells, and the cell where a vehicle exists is shown as "1", and the cell where a vehicle does not exist is shown as "0". An individual vehicle moves using the rule-184 [2]: advance one mass only when a forward cell empties. The updating time of the rule-184 is shown as,

$$U_j^{t+1} = U_j^t + \min(U_{j-1}^t, 1 - U_j^t) - \min(U_j^t, 1 - U_{j+1}^t) \qquad (1)$$

t: time , j: position of cell

U_j^t : state of position j at time t

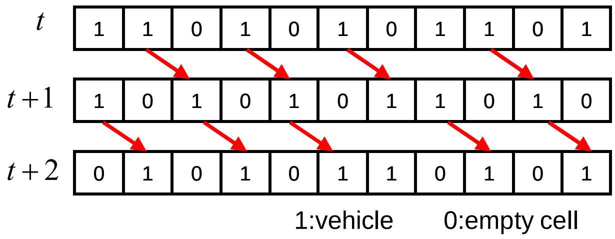

Fig.1 Movement of vehicles by rule-184

The movement of the vehicles by the rule-184 is shown in Fig.1. This model can be extended to a two-dimensional model which includes crossroads.

III. TRAFFIC SIGNAL CONTROL BY Q-LEARNING

The signal agent is composed of the State observer, the Learning, and the Select action as shown in Fig.2.

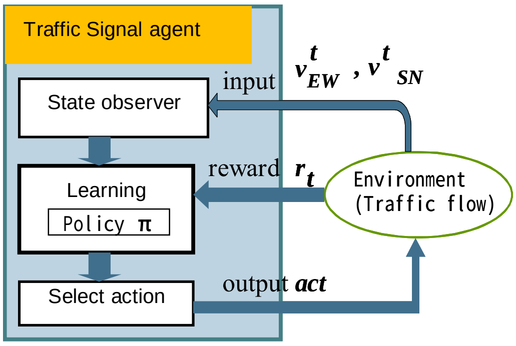

Fig.2 Structure of Traffic Signal agent

The State observer measures the number of the inflowing and outflowing vehicles from/to the crossroad and the stationary vehicles between adjoining crossroads.

The number of the stationary vehicles is denoted as follows:

v^t_{EW} : Vehicles existed on the road in the direction of east and west at time t

v^t_{SN} : Vehicles existed on the road in the direction of south and north at time t

The Learning carries out the update for the parameters of the traffic signal using the Q-learning which is an environmental identification type.

The action value function is defined as $Q(v^t_{EW}, v^t_{SN}, act)$.

act: control parameter which yields the passage in direction of east and west

Q value is updated every two cycles. When the number of the stationary vehicles between adjoining crossroads increases more than the number of the stationary vehicles of the previous cycle, the signal agent gets a penalty r, and the penalty is calculated for each inflowing road. Moreover, when the number of the stationary vehicles decreases, a reward is given. Therefore, the updating of Q value is defined as

$$Q(v^t_{EW}, v^t_{SN}, act_t) \leftarrow Q(v^t_{EW}, v^t_{SN}, act_t) + \alpha[r_t + \gamma \max_{a \in A} Q(v^{t+1}_{EW}, v^{t+1}_{SN}, act_{t+1}) - Q(v^t_{EW}, v^t_{SN}, act_t)] \quad (2)$$

α : learning rate $(0 < \alpha \leqq 1)$,

γ : discount rate $(0 \leqq \gamma < 1)$

The ε-greedy selection is used for the Select action. It acts randomly at the probability ε, and except for this situation, the agent acts according to the Q values memorized in the learning machine.

IV. SIMULATION

1. Field of Simulation and Q-learning Process

The overview of a road network consisted from three crossroads will be targeted in this simulation as shown in Fig.3. There is a road with two lanes including a right turn lane (6 cells, 43m) on Rikimaru, Kitakyushu-city. The distance between adjoining crossroads is 64 cells (480m). The number of the traffic density for the crossroads area was measured between 17:00 and 19:00 on January 17th, 2008. Based on the measured date, the following probabilities are supposed: the vehicles flow into the road network from the two ends of east-west direction and the six ends of north-south direction with probabilities 0.35 and 0.1, respectively; the vehicles flow outside of the road network with probability 0.8; the vehicles flow into the crossroads from the straight path, the left turn, and the right turn in the direction of north-south with probabilities 0.2, 0.7, and 0.1, respectively, and in the direction of east-west with probabilities 0.8, 0.15, and 0.05, respectively.

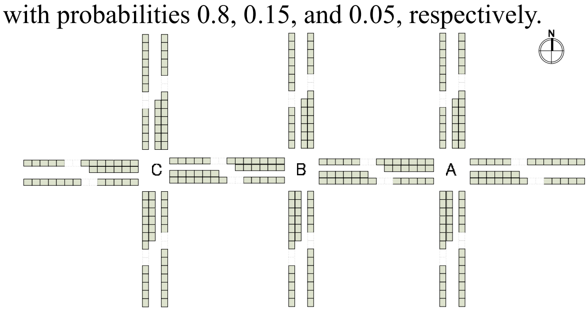

Fig.3 Gridiron

The splits in the signal of crossroads A, B, and C are assumed to be the same ratio. The behaviors of the signal agent are assumed to be the following:

Action = {10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%}

Fig.4 shows the procedure of the simulation by Q learning for the splits.

Initialize $Q(s, a)$ arbitrarily;
Repeat (for each episode);
  Repeat (for each step of episode);
    Observe a state observation *st* ;
    Choose *at* from *st* using policy derived from Q value;
    Take action *at*, observe *rt*, *st*+1;
    Q value is updated by the following update equation;

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\{r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)\}$$

$t \rightarrow t+1$
until *st* is terminal;
until all episodes are finished;
where α : Learning rate $(0 < \alpha \leqq 1)$
γ : Discount rate $(0 \leqq \gamma < 1)$

Fig.4 flow of procedure

The offset values of the crossroads A and B are defined based on the crossroad C. The amount of the change of the absolute offset value assumes 1/4 or less [3] of the cycle length 250(steps), and calculates each signal as the offset= {0, 20, 40, 60} (steps). Moreover, one cycle is 60 steps, and in the finish time of one split, all the traffic signals are red light during two steps.

2. Result of Simulation

The average number of the stationary vehicles between adjoining crossroads A and B is decreased about 1/4 for the east-west split 70% compared with the east-west splits 50% and 80% on 300,000 advanced steps as shown in Fig.5. Through the learning, the traffic signal agent can obtain the split to reduce the traffic jam.

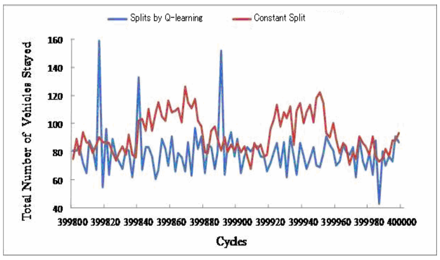

Fig.5 Total number of the vehicles stayed (Offset=0)
Average: 81.6 vehicles
Standard Variance: 12.9

The number of the east-west splits selected during the learning for the traffic signal agent on the crossroad A is shown in Fig.6, and the most selected split is 60%.

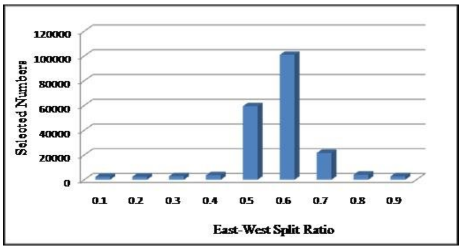

Fig.6 Selected numbers of East-West Sprit Ratio

Table1 Sprit

		Standing vehicles in the direction of North-South									
		10 ~19	20 ~29	30 ~39	40 ~49	50 ~59	60 ~69	70 ~79	80 ~89	90 ~99	100 ~109
Standing vehicles in the direction of East-West	10～19	50	50	50	50	40	50	50	50	50	50
	20～29	80	30	60	20	50	10	20	50	30	30
	30～39	70	60	70	80	40	60	50	30	40	40
	40～49	60	60	30	10	70	50	60	30	40	50
	50～59	60	60	60	30	80	50	50	40	40	40
	60～69	50	60	40	60	40	40	40	30	70	50
	70～79	50	60	50	30	30	80	40	50	50	50
	80～89	50	50	60	10	40	60	50	20	20	90
	90～99	50	60	40	50	50	50	50	30	70	60
	100～109	50	60	40	50	50	30	50	50	50	10
	110～119	60	60	50	50	50	30	40	40	60	80
	120～129	70	60	50	50	50	50	40	50	70	60
	130～139	60	60	60	50	50	40	50	80	10	90
	140～149	60	70	50	50	70	60	50	10	90	70
	150～159	60	60	50	50	70	50	50	80	40	80
	160～170	60	60	50	50	50	50	80	70	20	50

The splits of each stationary vehicle in the directions of east-west and north-south are shown in table1, and the total number of the stationary vehicles using the splits obtained by Q-learning and the constant split(50%) for each offset 20, 40, and 60 are shown in Fig.7, Fig.8, and Fig.9. As the result, the average of the stationary vehicles for the offset 20 is smaller than the other offsets.

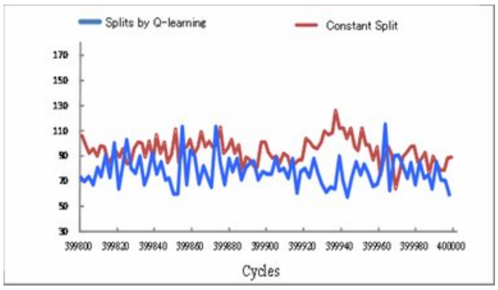

Fig.7 Total number of the vehicles stayed (Offset=20)
Average: 78.3 vehicles
Standard Variance: 13.6

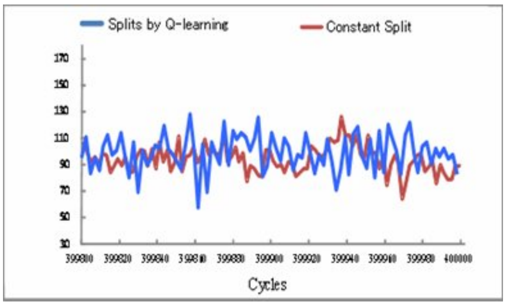

Fig.8 Total number of the vehicles stayed (Offset=40)
Average: 98.8 vehicles
Standard Variance: 12.8

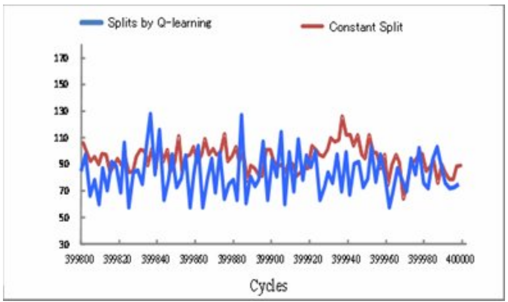

Fig.9 Total number of the vehicles stayed (Offset=60)
Average: 82.6 vehicles
Standard Variance: 14.8

REFERENCES

[1] S.Wolfrm, Theory and application of cellular automata, World Scientific, Singapore (1986)
[2] Naofumi Yamamoto, Shin Morishita, Simulation of Traffic Flow by Cellular Automata and its Control (In Japanese), Transactions of the Japan Society of Mechanical Engineers, Series C, Vol. 65, No. 637, pp.75-80 (1999)
[3] Satoshi Igaue, Mamoru Hisai, Real-Time Control of Coordinated Traffic Signal Systems through Sequential Formation of Offsets (In Japanese), Memoirs of the Faculty of Engineering, Yamaguchi University, Vol.56, No.1, pp.53-61 (2005)

