Traffic light control based on Fuzzy Q_learning

Marjan Jalali Moghaddam
Amirkabir University of Technology
Tehran, Iran
jalalymarjan@aut.ac.ir

Matin Hosseini
Amirkabir University of Technology
Tehran, Iran
matinhosseiny@aut.ac.ir

Reza Safabakhsh
Amirkabir University of Technology
Tehran, Iran
safa@aut.ac.ir

***Abstract*—Traffic is an issue that many big cities are confronted with because of ever-increasing population growth. In this paper we propose a two phase traffic light control system based on fuzzy Q-learning for an isolated 4-way intersection. The states and actions of the Q-learning variables is set by a fuzzy algorithm which can be learned through environmental interactions and taking advantage of fuzzy logic. The proposed algorithm was simulated for a period of one hour for each of 14 different traffic conditions. Comparison with other methods was carried out on the 14 traffic conditions. The results showed that the proposed algorithms decrease the total waiting time and the mean of queue length.**

***Index Terms*—Fuzzy logic, fuzzy Q-learning, traffic light control.**

I. Introduction

Many big cities are confronted with heavy traffic because of the ever-increasing population growth. Efficient methods for traffic flow management are necessary to optimize the use of available road capacity. Furthermore, high fuel prices and environmental problems are other important reasons to reduce traffic. Therefore, the need arise for simulating and optimizing traffic control system that can adapt with this increasing congestion. We tried to decrease traffic congestion through the development of an intelligent traffic control system, which is based on the traffic density flow at the intersection. Many research activities have attempted to solve different traffic problems in the intelligent transportation system [1,2,3,4,5]. In many cases, a fixed time traffic light control system has been used with the aim of minimizing waiting time and the number of vehicles waiting at intersections. However, it is of interest to apply dynamic traffic light control systems in which green light duration is adjusted to environmental dynamic changes to maximize throughput and minimize the waiting time. Fuzzy control systems use fuzzy logic, which can simulate human intelligence to control traffic and enable implementation of real world rules and human-like thinking process. Fuzzy control is an approach that can be applied to various traffic models. The fuzzy logic traffic light controllers utilize sensors to count the number of vehicles. Therefore, the resulting controllers control the traffic lights according to traffic density [6]. There are different approaches to control traffic lights based on artificial intelligence methods such as fuzzy logic, neural networks, reinforcement learning and evolutionary algorithms. These methods can lead to shorter queues and less traffic delay.

Many studies have focused on intelligent traffic light control. There are two approaches for traffic light control, including pre-timed [7] and sensor-based signal controls. **The preset cycle time** methods present the traditional strategy and consist of a preset interval series which cannot respond to unpredictable conditions and waste of time for less congested roads [8]. **The vehicle actuated** methods are sensor based methods in which the green light time can be extended based on demand.

To overcome the mentioned problems and in order to reduce the waiting time and the queue length, we proposed a Q-learning fuzzy controller which is based on the traffic density. As this paper considers the traffic light control approach, the existing methods in this category are briefly discussed below.

Smith et al. proposed a neural network based approach for traffic light control. This approach has a time consuming learning process and reduces waiting time by 10% [9]. The use of fuzzy logic provides a high quality performance approach to control the traffic congestion [10,11]. Arora et al. measured traffic density on the road using morphological edge detection and a fuzzy logic technique [12].

Tari et al. used a two level hierarchical fuzzy rule-base system for controlling complex traffic intersection [13]. Keyarsalan et al. used computer vision techniques and neural networks to extract the traffic data and applied a fuzzy ontology to control traffic light in isolated intersection [6]. Shakeri et al. introduced a three layer fuzzy system based on the cellular automata for optimizing traffic light control [14].

Abdulhai et al. provide an isolated traffic signal controller using reinforcement learning which could combine with dynamic route guidance [15]. Also multi-agent Q-learning was used for a non-stationary environment that estimated states based on the average queue length [16].

In order to minimizing the waiting time of the public transportation and reducing the computational complexity, dynamic programming and branch-and-bound techniques were combined to control traffic lights [17].

Liu et al. presented a differential evolution bacteria foraging optimization algorithm to minimize the delay vehicles of a cycle time, and also maximize throughput of the intersection [18]. In another work, vehicles were detected by the Prewitt edge detector and image matching and the traffic light duration was determined based on the percentage of matching [19].

Dujardin et al. applied mixed integer linear programming (MILP) for multimodal traffic light control based on optimization of three criteria including the total delay of persons and public vehicles, and the number of stops for private vehicles [20].

In this paper, we propose a two phases real-time approach based on fuzzy Q-learning for adjusting traffic light duration for an isolated intersection based on traffic flow. Compared to pre-set cycle time (PCT) and vehicle actuated (VA) approaches, the proposed method can reduce the average number of vehicles in traffic queues and the average waiting time for vehicles. The proposed system can learn the various relationships between traffic conditions and the optimal actions using its experience in different situations. Furthermore, it can effectively work in specific situations based on its past experience with the same or similar situations.

The advantages of the proposed approaches are as follows:

- No need to pre-specified models, and training possible for any traffic conditions.
- Can learn the relationships between states and actions using environmental interactions.
- Benefit from the fuzzy system advantages.

The rest of this paper is organized as follows. Section 2 reviews fuzzy Q_learning and presents the algorithm scheme. Section 3 reports experimental results and section 4 concludes the paper.

II. Proposed Method

The proposed method is shown in Fig. 1. Q-learning has been used for learning fuzzy systems [21]. Since the state and actions of Q-learning algorithm can be set by fuzzy variables, Q-learning can take advantage of fuzziness.

In this paper, a fuzzy Q-learning controller is designed for an isolated 4-way traffic intersection. The Q-learning information is used in tuning the output membership functions of the fuzzy controller. Before the end of each phase, the next green optimum phase durations estimated based on current traffic conditions which are specified by four variables include in the number of vehicles in the north, the south, the east and the west which are shown by ql_north, ql_south, ql_east, ql_west, respectively. Two input variables are considered for the traffic lights control:

- Max_ql_ns: the maximum number of vehicles in the north-south and the south-north (max (ql_north, ql_south))
- Max_ql_ew: the maximum number of vehicles in the east-west and west-east (max (ql_east, ql_west))

The green light and red light specified the arrival side and queue side, respectively. The proposed algorithm determined the optimal next phase duration for the queue side based on the current queue lengths. The north-south and east-west Q-tables are used where size is determined as:

$$Q_{table} - size = noas * noqs * NOaction \quad (1)$$

where noas is the number of the arrival side membership functions and noqs is the number of the queue side membership functions, and $NOaction$is the number of actions.

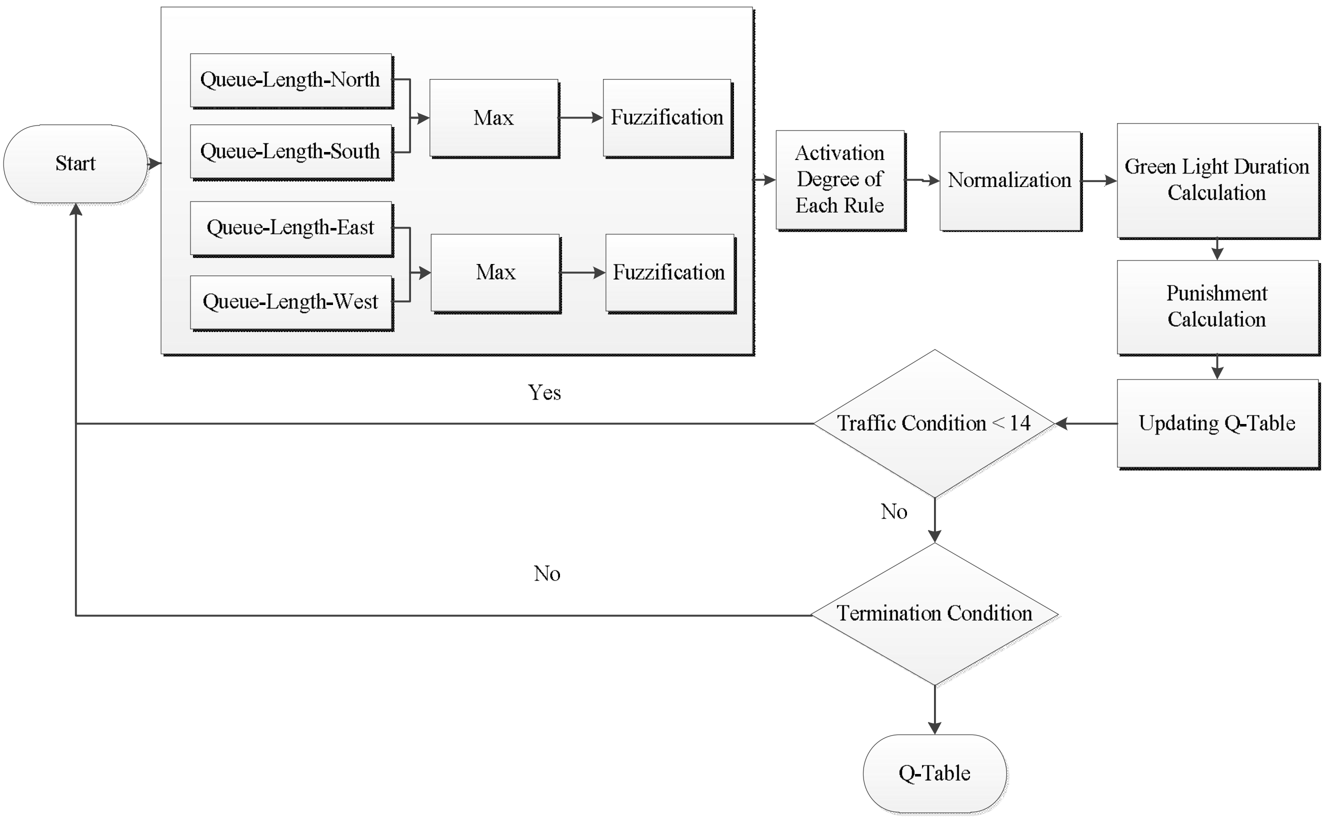

Figure 1. Fuzzy Q-Learning Traffic Light Control Schematic.

The proposed algorithm is described as follows:
1. Based on the number of vehicles in each queue, the variables Max_ql_ns and Max_ql_ew are calculated as an input for the next step of the algorithm.
2. Four fuzzy sets are defined on each dimension of the two dimensional state space, mf_{ns}^{i} and mf_{ew}^{i} are the corresponding membership function in which i ϵ {low, medium, high, very high} as are illustrated in Fig. 2. Each rule is associated with a set of possible discrete actions $Act_{j_1,j_2} = \{act_{j_1,j_2,r_1}, act_{j_1,j_2,r_2}, \ldots, act_{j_1,j_2,r_k}\}$

where j_1, j_2= 1, 2 ,… , show the number of membership functions, and r = 1, 2, …, n_R where n_R is the number of rules. The corresponding action values are defined as follow:

$$Q_{j_1,j_2} = \{q_{j_1,j_2,r_1}, q_{j_1,j_2,r_2}, \ldots, q_{j_1,j_2,r_k}\} \qquad (2)$$

According to the above definitions, the generic rule R_r may be written as follows:

R_r : If Max_ql_ns is mf_{ns}^{i} and Max_ql_ew is mf_{ew}^{i}

Then output = $a_{r,1}$ with $q_{r,1}$

OR
.
.
.
OR

output = a_{r,n_R} with q_{r,n_R}

When Max_ql_ns and Max_ql_ew enter the system, they are fuzzified based on the membership function. All of the rules are activated partially by a certain activation level which is calculated as follow:

$$\varphi_r = \mu_{ns}^{i}(\text{Max\_ql\_ns}) * \mu_{ew}^{i}(\text{Max\_ql\_ew}) \qquad (3)$$

where μ_{ns} and μ_{ew} are truth degrees.
3. Since each input variable belongs to several fuzzy sets with different activation levels, the activation level is normalized and is considered as the weight of each of the winning actions of rule:

$$\varphi\_\text{norm}_r = \frac{\varphi_r}{\sum_{r=1}^{N_r} \varphi_r} \qquad (4)$$

4. The time of green light is calculated based on the average weight method:

$$Act = \sum_{r=1}^{N_r} \varphi_r * win\_act_r \qquad (5)$$

where φ_r and win_act_r are the weight and wining action of the r[th] rule, respectively.

5. After the calculation of the next green light duration, the output value which is between -1 and 1, is mapped to the original range. In this paper, the original range is multiplies of 5 between 10 and 100.
6. After each phases, the punishment is calculated based on the queue length variation.

$$\text{Punishment} = \sum_{i=1}^{4} (\log(\max(|ql_i^{new} - ql_i^{old}|), 1) * sgn(ql_i^{new} - ql_i^{old})) \qquad (6)$$

where $ql_i \in \{ql_{north}, ql_{south}, ql_{east}, ql_{west}\}$
7. After calculating the punishment, the arrival side Q table is updated. Gradient descent is usually used to update the parameters of algorithm:

$$q_{r,i}^{t+1} = q_{r,i}^{t} - \alpha \varepsilon_Q^t \psi_r$$
$$\varepsilon_Q^t = \text{Punishment} - (\gamma \max_a Q(s^{t+1}, a) - Q^t(s^t, A^t(x^t))) \qquad (7)$$
$$\max_a Q(s^{t+1}, a) = \sum_{r=1}^{n_R} \psi_r \max_i \{q_{r,i}^t\}$$

where γ is discount factor and α is learning rate.
8. The above steps are done for each of the 14 different traffic conditions in one hour simulations and then are repeated until convergence.

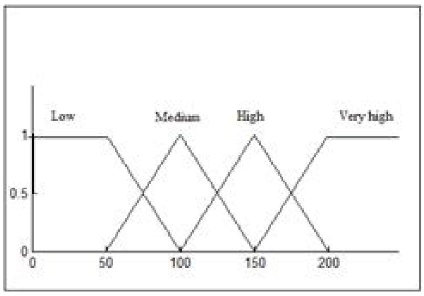

Figure 2.The membership functions for arrival side and queue side.

III. Experiments

A. Experimental Setting

The proposed algorithm for traffic light control is simulated for one hour for each of the 14 traffic conditions presented in table 1 [22]. Vehicle arrival and departure rates are simulated using Poisson distributions on each street as shown in table 1. For all conditions departure rates are set equal to 1. This algorithm is implemented using Matlab on an Intel® core i5 M460 2.53 GHz PC with 4 GB RAM. In order to evaluate the proposed algorithm, the parameter values are set as: γ =0.8, ε=0.01, and α =0.2. After each 14 states, alpha is updated with a 0.99 update rate. In each phase, based on the ε-greedy exploration strategy, the best ction with probability 1 − ε and a random action with probability ε is selected for each rule.

Table 1. Intersection traffic conditions [22]

Traffic condition	state	Input rate				description
		North	south	east	west	
balanced	I	0.25	0.25	0.25	0.25	Light traffic
	2	0.5	0.5	0.5	0.5	Moderate traffic
	3	0.75	0.75	0.75	0.75	Heavy traffic
Unbalanced	4	0.25	0.5	0.5	0.5	North has light traffic and south, east and west have moderate traffic
	5	0.5	0.25	0.25	0.25	North has moderate traffic and south, east and west have light traffic
	6	0.5	0.5	0.25	0.5	East has light traffic and north, south and west have moderate traffic
	7	0.5	0.5	0.5	0.25	west has light traffic and north, south and east have moderate traffic
	8	0.75	0.25	0.25	0.25	North has heavy traffic and south, east and west have light traffic
	9	0.25	0.75	0.25	0.25	South has heavy traffic and north, east and west have light traffic
	10	0.25	0.25	0.75	0.25	east has heavy traffic and north, south and west have light traffic
	11	0.25	0.25	0.25	0.75	west has heavy traffic and north, south and east have light traffic
complementary	12	0.25	0.5	0.25	0.5	North and east have light traffic and south and west have moderate traffic
	13	0.75	0.5	0.75	0.5	North and east have heavy traffic and south and west have moderate traffic
	14	0.25	0.75	0.25	0.75	North and east have light traffic and south and west have heavy traffic

Table 2. The mean of total waiting and mean of q length of diffrent algorithms.

Traffic Condition	Algorithm					
	Fuzzy Q-Learning		PCT		VA	
	MQL	MTW	MQL	MTW	MQL	MTW
1	1.71	7.81	1.8	6.74	1.46	5.49
2	29.81	69.5	26.53	55.05	34.28	78.72
3	293.84	496.27	447.85	603.52	397.09	592.21
4	34.24	71.71	28.01	57.56	29.91	84.06
5	28.54	60.14	27.75	57.64	24.08	64.45
6	21.34	45.17	22.91	47.03	22.42	51.97
7	21.78	47.31	26.7	54.64	25.62	63.46
8	27.85	66.66	121.67	164.6	84.22	135.28
9	29.30	77.77	105.07	145.61	85.16	136.3
10	24.03	49.67	103.36	145.82	81.07	143.28
11	33.19	86.47	107.43	145.69	72.42	117.97
12	1.29	7.05	1.09	5.26	1.1	5.43
13	229.71	367.58	246.06	349.97	237.38	367.6
14	38.84	74.85	236.6	328.79	102.59	168.06

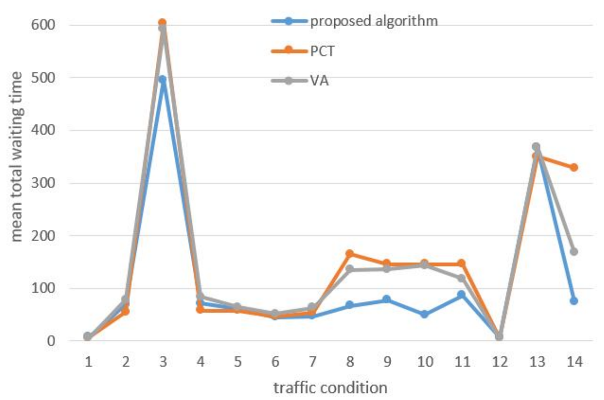

Figure 3. The mean total waiting time

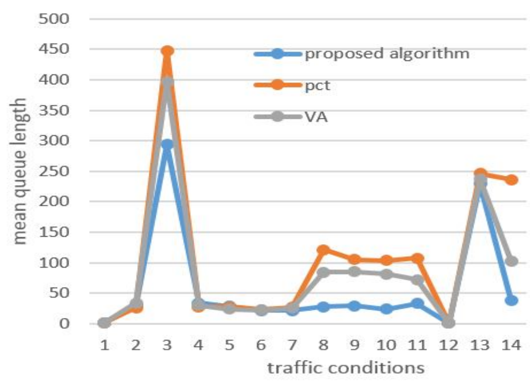

Figure 4. The mean of queue length

B. Experimental Results

In this section, we present the results of our experiments. We assessed and compared the proposed algorithm with two other traffic light control methods, namely PCT and VA on the 14 traffic conditions. Based on the arrival and departure time of each vehicle, the total waiting time and mean queue length are calculated.

The mean of total waiting time (MTW) and mean of queue lengths (MQL) are calculated for each algorithm in table 2.

According to table 2, the mean of queue length and mean of total waiting time of the results are nearly similar in the light, and moderate traffic condition which are colored in black. The important case appears in the heavy traffic conditions. Our algorithm achieves a better result than the other algorithms, especially in unbalanced traffic conditions, where the input rate of one side is heavy and other methods show critical results.

The mean total waiting time of the three algorithms are compared in Fig. 3. PCT and VA face difficulties in the more unbalanced condition and heavier input rate. However the proposed algorithm achieves better results and can deal with such conditions. The comparison of the mean of queue length of three algorithms is illustrated in Fig. 4.

In asymmetrical traffic increase from different directions, the fuzzy Q-learning controls traffic in shorter mean queue length and total waiting time.

IV. Conclusions

We proposed a two phase traffic light control system based on fuzzy Q-learning for an isolated 4-way intersection. A fuzzy algorithm set the Q-learning variables. The proposed algorithm benefits from fuzzy system advantages and can learn the fuzzy rules using environmental interactions. Before the end of each phase, the next green optimum phase durations estimated based on current traffic conditions which are specified. This algorithm was compared with two other algorithms, namely VA and PCT for a period of one hour for each of 14 different traffic conditions. As shown in the results, the proposed algorithm outperformed the other algorithms significantly in heavy traffic and unbalanced traffic conditions, whereas in other situations the three algorithms were similar.

REFERENCES

[1] Jalali Moghaddam, M. Shabani, E., and Safabakhsh, R., “Crowd Density Estimation for Outdoor nvironments”,Proceeding of 8th International Conference on Bio-inspired Information and Communication Technologies, 2014, Boston, USA.
[2] Ghasemi, A. and Safabakhsh, R., “A real-time multiple vehicle classification and tracking system with occlusion handling,” IEEE 8th International Conference on Intelligent Computer Communication and Processing, Sep. 30, 2012, Romania.
[3] Ghasemi, A. and Safabakhsh, R., " .Unsupervised Foreground-Background Segmentation Using Growing Self-Organizing Map in Noisy Backgrounds,” 3rd International Conference on Computer Research and Development, 11-13 March, 2010, Shanghai, China.
[4] Kalaki, A. S., & Safabakhsh, R., (2014, February). Current and adjacent lanes detection for an autonomous vehicle to facilitate obstacle avoidance using a monocular camera. In Intelligent Systems (ICIS), 2014 Iranian Conference on (pp. 1-6). IEEE.
[5] Kalaki, A.S., and Safabakhsh, R., “Vision based Real-time Lane and Obstacle Detection and Tracking in Intelligent Vehicles,” 13th International Conference on Traffic and Transportation Engineering, Feb. 25-26, 2014, Tehran, Iran.
[6]Keyarsalan, M., and Gholam, A., "Designing an intelligent ontological system for traffic light control in isolated intersections." Engineering Applications of Artificial Intelligence 24.8 (2011): 1328-1339.
[7] WEBSTE COBBE, 1996,WEBSTER F. V., COBBE B. M. (1996). Technical Paper 56: Traffic Signals.
[8] Dotoli, Mariagrazia, Maria Pia Fanti, and Carlo Meloni. "A signal timing plan formulation for urban traffic control." Control Engineering Practice 14.11 (2006): 1297-1311.
[9] Smith, Richard H., and Daniel C. Chin. "Evaluation of an adaptive traffic control technique with underlying system changes." Proceedings of the 27th conference on Winter simulation. IEEE Computer Society, 1995.
[10]Hoyer, R., & Jumar, U. (1994, June). Fuzzy control of traffic lights. In Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on (pp. 1526-1531). IEEE.
[11]Hejun, W., & Changyun, M., (2010). Design of intelligent traffic light control system based on traffic flow. In 2010 International Conference on Computer and Communication Technologies in Agriculture Engineering (Vol. 3, pp. 368-371).
[12] Arora, Madhavi, and V. K. Banga. "Intelligent Traffic Light Control System using Morphological Edge Detection and Fuzzy Logic."
[13]Tari, T., Kóczy, L. T., Gáspár, C., & Hontvári, J. (2006). Control of Traffic Lights in High Complexity Intersections Using Hierarchical Interpolative Fuzzy Methods. In Fuzzy Systems, 2006 IEEE International Conference on (pp. 1045-1048). IEEE.
[14] Shakeri, Moein, et al. "A novel fuzzy method to traffic light control based on unidirectional selective cellular automata for urban traffic." Computer and Information Technology, 2008. ICCIT 2008. 11th International Conference on. IEEE, 2008.
[15] Abdulhai, Baher, Rob Pringle, and Grigoris J. Karakoulas. "Reinforcement learning for true adaptive traffic signal control." Journal of Transportation Engineering 129.3 (2003): 278-285.
[16] Abdoos, Monireh, Nasser Mozayani, and Ana LC Bazzan. "Traffic light control in non-stationary environments based on multi agent Q-learning." Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on. IEEE, 2011.
[17] Riedel, Th, and U. Brunner. "Traffic control using graph theory." Control Engineering Practice 2.3 (1994): 397-404.
[18] Liu, Qin, and Jianmin Xu. "Traffic Signal Timing Optimization for Isolated Intersections Based on Differential Evolution Bacteria Foraging Algorithm."Procedia-Social and Behavioral Sciences 43 (2012): 210-215.
[19] Choudekar, P., Banerjee, S., &Muju, M. K. (2011, April). Implementation of image processing in real time traffic light control. In Electronics Computer Technology (ICECT), 2011 3rd International Conference on (Vol. 2, pp. 94-98). IEEE.
[20] Dujardin, Y., Boillot, F., Vanderpooten, D., &Vinant, P. (2011, October). Multiobjective and multimodal adaptive traffic light control on single junctions. InIntelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on (pp. 1361-1368). IEEE.
[21] Bonarini, A., Lazaric, A., Montrone, F., & Restelli, M. (2009). Reinforcement distribution in fuzzy Q-learning. Fuzzy sets and systems, 160(10), 1420-1443.
[22] Chong, Y., Quek, C., & Loh, P. (2009). A novel neuro-cognitive approach to modeling traffic control and flow based on fuzzy neural techniques. Expert Systems with Applications, 36(3), 4788-4803.