OVERVIEW SYNTHESIS OF TRAFFIC CONTROL STRATEGIES – WHERE ARE WE NOW?

Daniela Koltovska[1]

[1]Department of Traffic and Transport, Faculty of Technical Sciences, St. Kliment Ohridski University, Republic of Macedonia,
daniela.koltovska@tfb.uklo.edu.mk

OVERVIEW OF TRAFFIC CONTROL STRATEGIES

Urban vehicular traffic is an expression of human behaviour and as such, it is highly variable in space and time. Therefore, a high degree of adaptiveness is required in traffic management systems for handling these variable traffic conditions with reasonable efficiency
Gartner, 1982

Table 1. Types of generations of adaptive control strategies

Feature	First Generation	Second Generation	Third Generation
Frequency of update interval	15 min	5 -10 min	3-5 min
Control plan generation	Off line optimization selection from a library by time of day, traffic responsive, or manual mode	Online optimization	Online optimization
Traffic prediction	No	Yes	Yes
Cycle length determination	Fixed within each intersection	Fixed within variable groups of intersections	Variable in time and space. Predetermined for control period

Source: Gartner, N., Stamatindius, C., Tarnoff, P., Development of Advanced Traffic Signal Control Strategies for ITS, Transportation Research Record 1494, 1996.

WHERE ARE WE TODAY?

*New levels, generations of traffic control - 4th and 5th generations of control. Level 5, it should be a "**superlevel**" that uses **Artificial intelligence (AI)** to select from among an array of existing control strategies (taken from Levels 1 through 4 based on traffic conditions.*
Abdulhai, 2003

Q – learning approach can achieve the functionality and benefits of Level 5GC "superlevel" becasue:

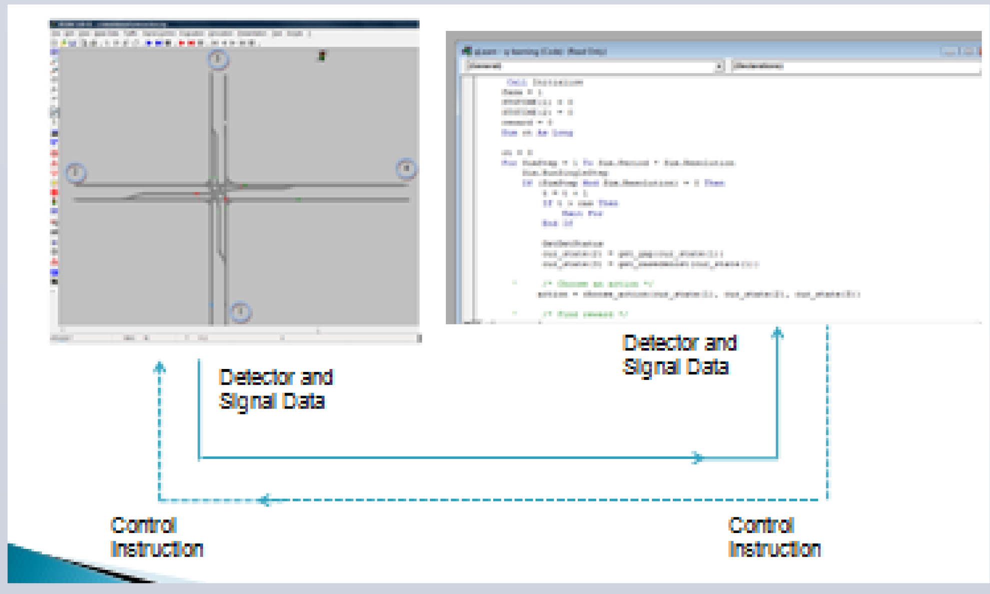
- *The ability to learn the best control decision for a given situation through a prior training period that involves an exhaustive trial-and-error process on a simulator followed by fine-tuning and automatic updating through interaction with real-world conditions.*
- *The ability to choose, in real-time, the best objective to pursue for a given situation. This may be a local objective such as reducing local delay, or a global objective such as maximizing the rate of traffic exiting an area.*
- *The ability to perform self-evaluation and undertake additional training (in the background) when it is recognized that the environment has changed or control decisions are no longer optimal.*
- *The ability to make use of emerging detection technology.*
- *The ability to do the above with a high degree of automation and seamlessness.*

Abdulhai, 2003

OWN APPROACH AND EXPERIENCE IN DEVELOPMENT OF ADAPTIVE TRAFFIC CONTROL STRATEGY

Development of traffic control strategy on isolated intersection with technique of reinforcement learning (RL) and the Q-learning approach

- *Defining the traffic network and detectors, creating the traffic demand*
- *Designing the operation of traffic signals*
- *Defining the agent of RL (state, actions, reward)*
- *Algorithm parameters testing process*



DEFINITION OF Q –LEARNING AGENT

- **Defining the states (S)**
- **Defining the actions (A):**
- Continue – traffic signals stay unchanged
- Change traffic signals
- **Defining the rewards (R)**
- The goal function is to maximize the vehicle throughput
- **Reward function** is the total throughput.
- Immediate reward is the number of vehicles passing when green **within the previous time interval** (duration is 90 seconds)
- Discounted rewar**d** is the total number of vehicles in the observed period of 3600 seconds (T – 3600 seconds, peak hour for which the strategy is tested)

The Q-learning algorithm (control algorithm of the temporary difference (off-policy TD Control)) is shown in its procedural form:

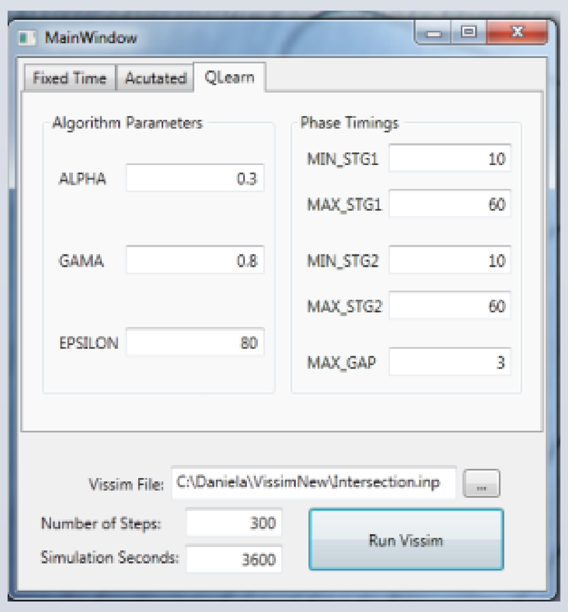
Initialize $Q(s,a)$ arbitrarily
Repeat (for each episode):
Initialize s
Repeat (for each step of episode):
Choose a from s using policy derived from Q (e.g., ε-greedy)
Take action a, observe r, s'
$Q(s,a) \longleftarrow Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$
$S \leftarrow s'$;
until s is terminal.

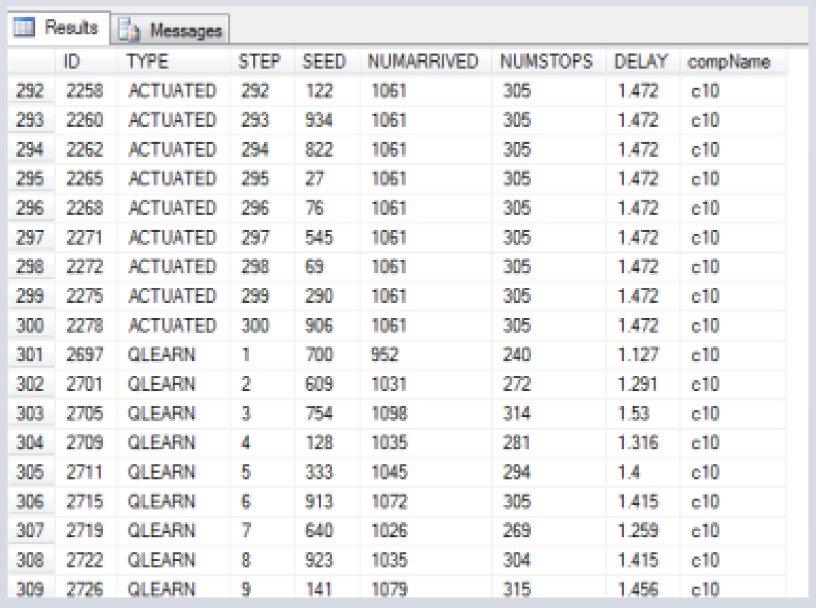
PROCESS OF THE ALGORITHM PARAMETERS TESTING

- The micro-simulator PTV VISSIM 5.4 and COM was used in the development of the strategy
- c# was used for the program part, whereas the data was collected in an ACCESS database and Microsoft SQL server

  The testing process with duration of 3600 seconds was performed in 300 steps
- In order to prove the convergence of Q values, different parameters of the algorithm were set

Algorithm Parameters	Testing Values
Learningrate α	0.9; 0.6; 0.3
Discount rate γ	0.5
Greedy policy ε	80; 40





Results | Messages

	ID	TYPE	STEP	SEED	NUMARRIVED	NUMSTOPS	DELAY	compName
292	2258	ACTUATED	292	122	1061	305	1.472	c10
293	2260	ACTUATED	293	934	1061	305	1.472	c10
294	2262	ACTUATED	294	822	1061	305	1.472	c10
295	2265	ACTUATED	295	27	1061	305	1.472	c10
296	2268	ACTUATED	296	76	1061	305	1.472	c10
297	2271	ACTUATED	297	545	1061	305	1.472	c10
298	2272	ACTUATED	298	69	1061	305	1.472	c10
299	2275	ACTUATED	299	290	1061	305	1.472	c10
300	2278	ACTUATED	300	906	1061	305	1.472	c10
301	2697	QLEARN	1	700	952	240	1.127	c10
302	2701	QLEARN	2	609	1031	272	1.291	c10
303	2705	QLEARN	3	754	1098	314	1.53	c10
304	2709	QLEARN	4	128	1035	281	1.316	c10
305	2711	QLEARN	5	333	1045	294	1.4	c10
306	2715	QLEARN	6	913	1072	305	1.415	c10
307	2719	QLEARN	7	640	1026	269	1.259	c10
308	2722	QLEARN	8	923	1035	304	1.415	c10
309	2726	QLEARN	9	141	1079	315	1.456	c10

FUTURE STEPS TO BE TAKEN

Different scenarios will be examined in the next stage for the purpose of a more realistic presentation of the traffic flows' stochastic nature as follows:

- Strategy testing at different arrival rates
- Strategy testing under conditions of different traffic demand flows on major and minor intersection approaches
- Testing under congestion
- This might involve certain modifications in the very control strategy