

Traffic Control and Optimization in ITS:

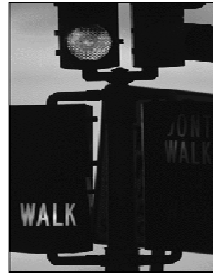
PART III.

Surface Street Control



Baher Abdulhai, Ph.D

ITS Centre and Testbed
University of Toronto



Overview

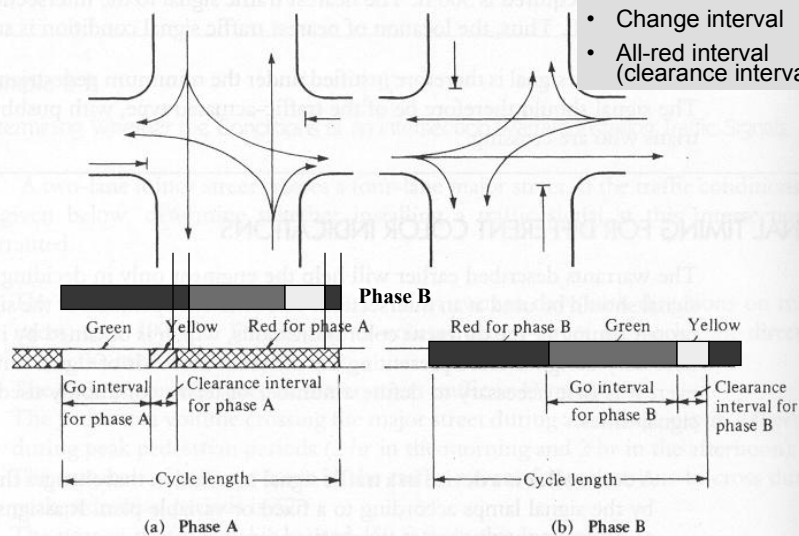
- Pre-timed signals
- Actuated signals
- Coordinated signals
- Dynamic signal control
- Adaptive signal control

Pre-timed Signals

- Terms relating to signal timing
- Objectives of signal timing
- Concept behind the determination of yellow interval & dilemma zone
- Determining cycle length for a pretimed, isolated signal using the Webster method
- Allocating green time to each signal phase

Terms relating to signal timing

- Cycle length
- Phase
- Interval
- Change interval
- All-red interval (clearance interval)

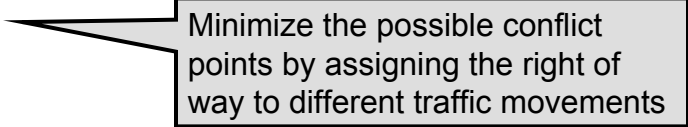


Terms relating to signal timing (cont)

- All signal analyses are based on through passenger cars.
- Conversion of:
 - other-than-passenger cars and
 - other-than-through vehicles
- Hence we need
 - Passenger Car Equivalent (PCU's)
 - Turning Movement Factor

Objectives of signal timing

- Reduce the average delay of all vehicles
- Reduce the probability of accidents



Minimize the possible conflict points by assigning the right of way to different traffic movements

Two conflicting objectives:

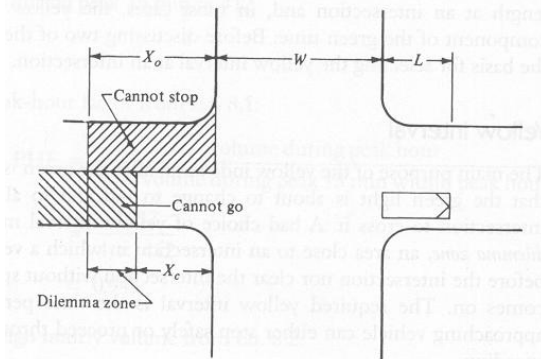
- More phases, less conflict
- More phases, more lost time



So, if at all possible, use:

- less phases
- Shorter cycle length

Yellow interval & Dilemma zone



$$X_0 = u_0 \delta + \frac{u_0^2}{2a}$$

X_0 : the min. distance from the intersection for which a vehicle traveling at the speed limit u_0 during the yellow interval cannot go through the intersection without accelerating. So it can stop.

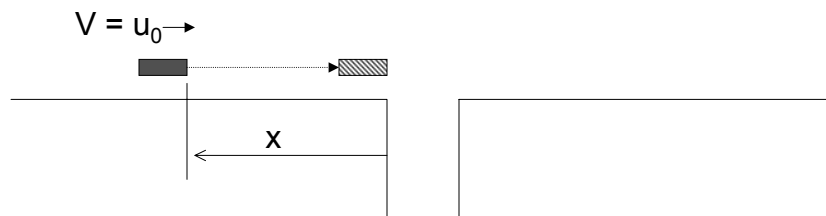
Distance traveled during yellow interval

$$u_0 (\tau_{\min}) = X_c - (W + L) \implies X_c = u_0 (\tau_{\min}) - (W + L)$$

X_c : the distance within which a vehicle traveling at the speed limit (u_0) during the yellow interval cannot stop before encroaching on the intersection.

Yellow interval & Dilemma zone (cont)

Safe stopping during Clearance interval



To stop $v = 0 = u_0 - at$

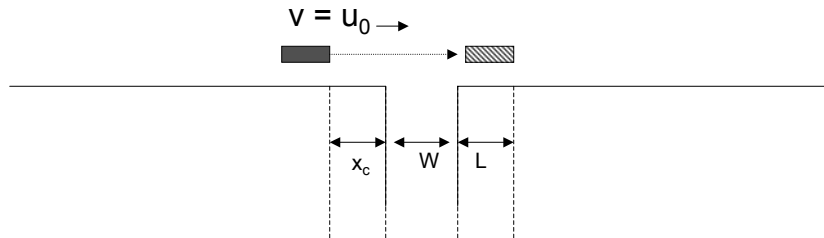
$T = \delta + u_0/a$

$X_0 = \delta u_0 + u_0^2/2a$

δ : perception time and reaction time
 a : deceleration rate

Yellow interval & Dilemma zone (cont)

Time needed to pass through the intersection during Clearance interval:



$$\tau_{\min} = (x_c + W + L) / u_0$$

$$x_c = \tau_{\min} u_0 - (W + L)$$

Yellow interval & Dilemma zone (cont)

The solution is to increase the time allowed to avoid $x_c < x_0$.

For the dilemma zones to be eliminated, $X_0 = X_c$:

$$u_0 \tau_{\min} - (W + L) = u_0 \delta + \frac{u_0^2}{2a}$$

δ Perception and reaction time

Solve for the yellow interval:

$$\tau_{\min} = \delta + \frac{u_0}{2a} + \frac{(W + L)}{u_0}$$

Comfortable deceleration rate:

$a = 0.27g$ (70% of g , where g is the gravitational acceleration)

If the effect of grade is added:

$$\tau_{\min} = \delta + \frac{u_0}{2(a + Gg)} + \frac{(W + L)}{u_0}$$

G = grade in decimals

Yellow = 3 – 5 sec: If longer, supplement with all-red

Yellow Computation Example

- Determine the minimum Yellow interval at an intersection whose:
 - width = 40 ft
 - Maximum allowable speed on the approach roads is 30 mph
 - Average length of vehicle = 20 ft

■ Solution:

$$\tau_{\min} = \delta + \frac{u_0}{2(a + Gg)} + \frac{(W + L)}{u_0}$$

$$= 1.0 + \frac{40 + 20}{30 \times 1.47} + \frac{30 \times 1.47}{2 \times 0.27 \times 32.2}$$

$$\Rightarrow \tau_{\min} = 4.9 \text{ sec} \Rightarrow \text{yellow period} = 5 \text{ sec}$$

Cycle lengths of pretimed signals by the Webster Method

$$C_0 = \frac{1.5L + 5}{1 - \sum_{i=1}^{\phi} Y_i}$$

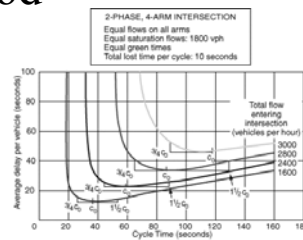


Figure D-6. Effect on delay of variation of the cycle length

L = Total lost time per cycle, usually you lose 3.5 seconds per distinct phases.

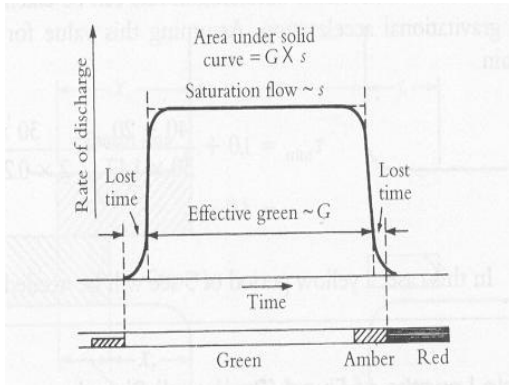
Y_i = Max value of the ratios of approach flows (called "**critical movement**") to saturation flows for all traffic streams using phase i , V_{ij}/S_j

ϕ = Number of phases

V_{ij} = Flow on lane j having the right of way during phase i

S_j = Saturation flow in one lane j , for a through lane, about 1900 pcphgpl

Cycle lengths of pretimed signals by the Webster Method (cont)



Lost time for phase i :

$$l_i = G_{ai} + \tau_i - G_{ei}$$

l_i = lost time for phase i

G_{ai} = actual green time for phase i (not including yellow time)

τ_{min} = yellow time for phase i

G_{ei} = effective green time for phase i

Total lost time:

$$L = \sum_{i=1}^{\phi} l_i + R \quad R = \text{Total all-red during the cycle}$$

Cycle lengths of pretimed signals by the Webster Method (cont)

Allocation of Green Times:

- In general the total effective green time available per cycle is given by:

$$G_{te} = C - L = C - \left(\sum_{i=1}^{\phi} l_i + R \right)$$

- Where C = actual cycle length used (usually obtained by rounding off C_0 to the nearest 5 sec)
- C_{te} = total effective green time per cycle
- To obtain minimum overall delay, the total effective green time should be distributed among the different phases in proportion to their Y values to obtain the effective green time for each phase:

$$G_{ei} = \frac{Y_i}{Y_1 + Y_2 + \dots + Y_{\phi}} G_{te}$$

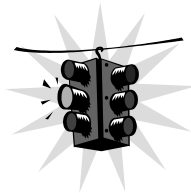
Cycle lengths of pretimed signals by the Webster Method (cont)

Allocation of Green Times:

- The actual green time for each phase is obtained as:

$$\left. \begin{aligned} \blacklozenge G_{a1} &= G_{e1} + l_1 - \tau_1 \\ \blacklozenge G_{a2} &= G_{e2} + l_2 - \tau_2 \\ \blacklozenge G_{ai} &= G_{ei} + l_i - \tau_i \\ \blacklozenge G_{a\phi} &= G_{e\phi} + l_\phi - \tau_\phi \end{aligned} \right\}$$

Usually l_i is about 3.5 sec.



Actuated Signals

Pretimed and Actuated Signals

- Pretimed signals operate with constant cycle lengths, phase sequences and interval timings
- But traffic demand varies with time that can cause isolated signal cycle, or series of cycles to fail
- Actuated signal attempts to address this type of problem through use of detectors on all or some of the approaches to the intersection

■ Semi-Actuated Control

- Detection only on minor side-street approaches to the intersection
- Green remains on main street until a “call” for service on side street is registered
- Frequently used to allow safe crossing of relatively light side-street traffic using signalization

■ Full-Actuated Control

- All the approaches have detectors
- Phases sequenced according to “calls” for service on all approaches
- Used at busy but isolated intersections, where demand fluctuations during the day make fixed-timing inefficient

Detection for Actuated Signalization

■ Primarily inductive loops are used for detection

➤ Point Detection

A single detector is placed for each approach lane

➤ Area Detection

Works with a long induction loop or a series of point detectors

➤ Modern video technology is rapidly developing

Advantageous as pavement need not be disturbed

Actuated Control Features and Operation

- The following features must be set on the controller
 - Minimum Green Time
 - Passage Time Interval
 - Maximum green Time
 - Recall Switch
 - Yellow Change Interval and All Red Clearance Interval

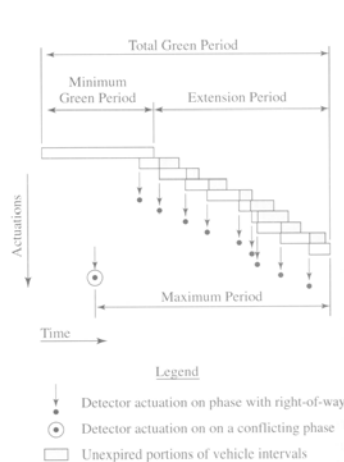
Actuated Control Features and Operation(contd.)

- Min Green Time:
 - Intends to provide sufficient time for vehicles stored b/w detector and stop line to enter the intersection
- Passage Time Interval
 - Allows a vehicle to travel from the detector to stop line
 - Also defines the maximum gap between vehicles arriving at the detector to retain a given green phase
- Maximum green Time
 - A green phase is retained to this limit if sufficient demand is there
 - After this time green terminates assuming a “call” on another phase

Actuated Control Features and Operation(contd.)

- Recall Switch
 - Each phase has one
 - When “on”, green is recalled from a terminating phase
 - When “off”, green is retained on previous phase until a “call” is received
- Yellow Change Interval and All Red Clearance Interval
 - Set for a fixed time as in pretimed controller

Actuated Control Features and Operation(contd.)



- Green retains at least for a specified min. green
- If a vehicle arrives during min. green, and less than one passage time left before termination, green time equal to passage time added from time of actuation
- Green terminated if
 - Passage time elapses without actuation
 - Max. green reached and “call” from another phase

Figure: Operation of an actuated phase

Signal Timing Parameters and Detector Placement

■ Minimum Green Time

$$G_{\min} = 4 + \left[2 \times \text{Integer} \left(\frac{d}{20} \right) \right]$$

$G_{\min}$ = initial interval(sec)
 d = distance b/w detector and stop line, ft
 20 = assumed distance between stored vehicles, ft
 4 = assumed start-up time, sec
 2 = assumed saturation headway, sec

■ Passage Time and the Allowable Gap

$$P = \frac{d}{1.468S}$$

P = passage time, sec
 d = distance b/w detector and stop line, ft
 S = approach speed of vehicles, mph

Signal Timing Parameters and Detector Placement

■ Maximum Green Times

- Set by working out an optimal cycle length and phase splits as if the controllers were pretimed

■ Some Detector location Strategies

- Using Equation for $G_{\min}$,

$$d = 10 G_{\min} - 40$$

- Setting detector back at a distance traveled at 3 to 4 sec

$$d = 1.468S(3) = 4.404S \text{ or } d = 1.468S(4) = 5.872S$$

Recommended Detector Locations and Timing Parameters

Approach Speed		Detector Set-Back (To front of loop)		Minimum Green	Passage Time
mph	kph	ft	m	sec	sec
15	24	40	12	8.0	3.0
20	32	60	18	10.0	3.0
25	40	80	24	12.0	3.0
30	48	100	30	14.0	3.5
35	56	135	41	18.0	3.5
40	64	170	52	22.0	3.5
45+	72+	Volume-density or multiple detectors recommended.			

Note: Volume-density could be considered at speeds of 35 mph (56 kph) or above.

Signal Coordination

Signal Coordination

Objective

- Coordinate green times of relatively close spaced signals to move vehicles efficiently through the set of signals
- Vehicles released from a signal maintain their grouping for more than 1000 ft.
- So, common practice to coordinate signals less than one-half mile apart on major streets and highways

Signal Coordination

Some Points

- All signals have the same cycle length
- Signal offset
 - Difference between two green initiation times
 - In general, measured in terms of d/s green initiation relative to u/s green initiation

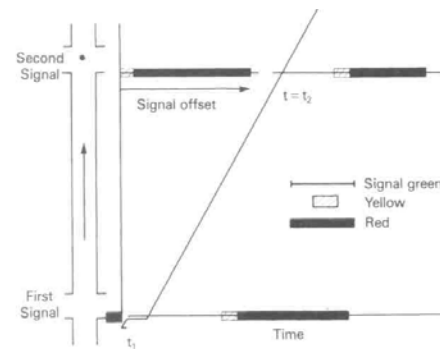


Figure: Illustrative vehicle trajectory

■ Benefits

- Considered in terms of a cost or penalty function

Cost = A*(total stops) + B*(other terms)

Objective to minimize this disbenefit
A, B are specified coefficients

- Reduces stops and delay
- Maintains a preferred speed

■ Factors lessening benefits

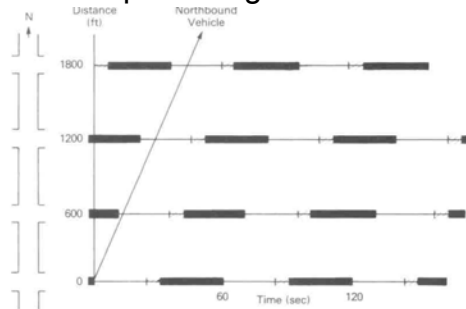
- Inadequate roadway capacity
- Road side frictions
- Complicated intersections
- Variability of traffic speed
- Very short signal spacing

■ Exceptions to the coordinated scheme

- Critical intersection

Time-Space Diagram and Ideal Offsets

■ Time-space diagram



Plot of signal indications as function of time for two or more signals

■ Ideal offset

➤ $t(\text{ideal}) = L/S$

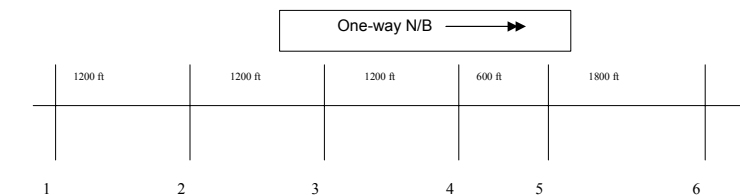
$t(\text{ideal}) = \text{ideal offset, sec}$

$L = \text{block length, ft}$

$S = \text{vehicle speed, fps}$

Signal Progression on One-way Streets: A Case Study

■ Determining ideal offsets



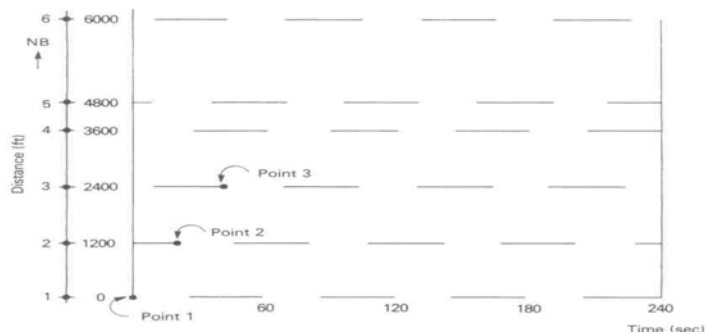
Intersection
All splits 50/50
 $C = 60 \text{ sec}$

Signal	Relative to Signal	Ideal Offset
6	5	$1800/60 = 30 \text{ sec}$
5	4	$600/60 = 10 \text{ sec}$
4	3	$1200/60 = 20 \text{ sec}$
3	2	$1200/60 = 20 \text{ sec}$
2	1	$1200/60 = 20 \text{ sec}$

Assuming a desired platoon speed of 60 fps

Signal Progression on One-way Streets : A Case Study(contd)

■ Time-space diagram



Signal Progression on One-way Streets

■ Effects of Vehicles queued at Signals

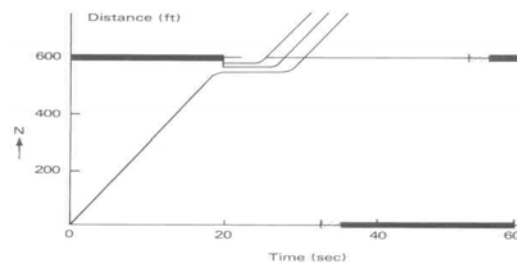
- Ideal offset must be adjusted for these vehicles

$$t_{ideal} = \frac{L}{S} - (Qh + Loss_1)$$

Q= # of vehicles queued per lane, veh

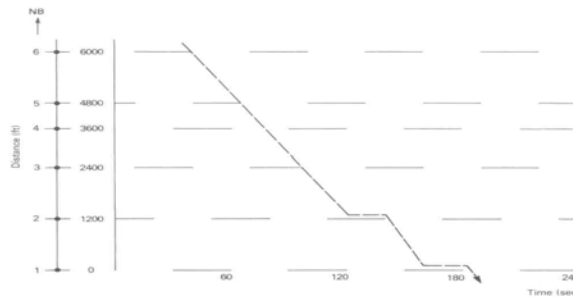
H=discharge headway of queued vehicle,sec/veh

Loss₁= initial loss time



Signal Progression on Two-way Streets

■ Offset determination on a two-way street



Moving s/b on a two-way arterial set for n/b progression:
An illustration

Signal Progression on Two-way Streets(contd)

■ Offset determination on a two-way street

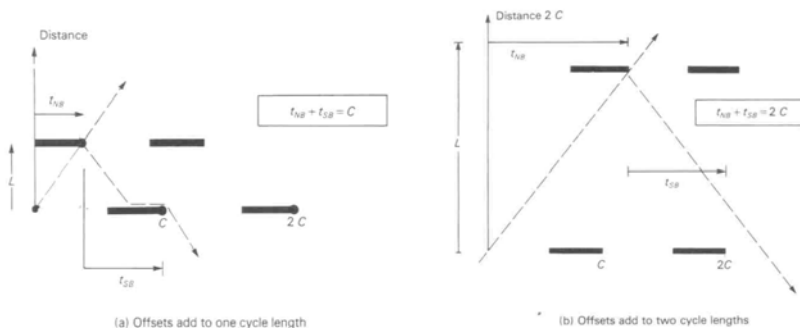


Figure: Offsets on a two-way arterial are not independent

Signal Progression on Two-way Streets(contd)

■ Offset determination in a grid

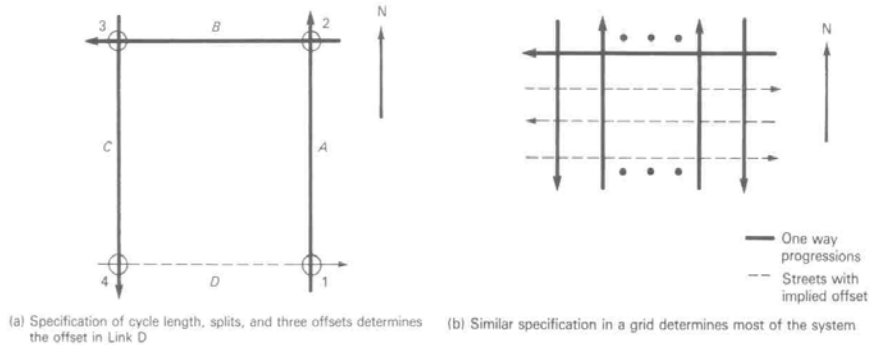


Figure: Illustration of the "closure" effect in a grid

Signal Progression on Two-way Streets(contd)

■ Offset determination in a grid

$$nC = t_A + g_{NS,2} + t_B + g_{EW,3} + t_C + g_{NS,4} + t_D + g_{EW,1} \dots (1)$$

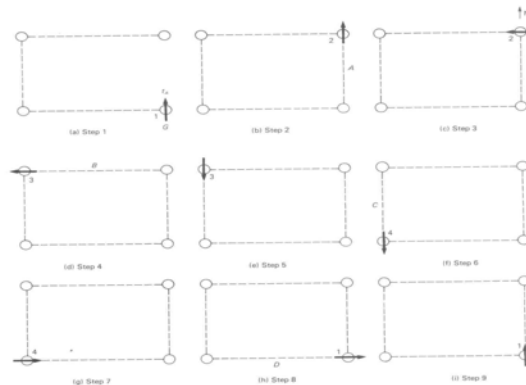


Figure: Development of Equation (1) for closure constraint in a network

Bandwidth Concept and Maximum Bandwidth

- “Windows” of green through which platoons of vehicles can move
- Efficiency of a progression = $\text{bandwidth} / \text{cycle length} \times 100\%$
 - An efficiency of 40-55% is considered good

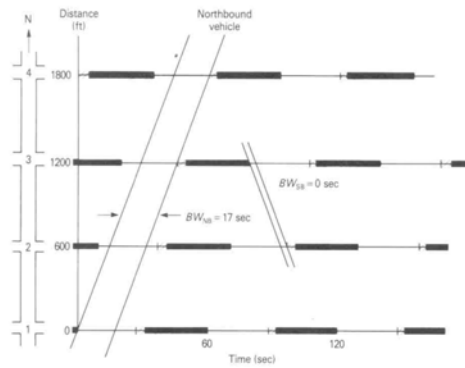


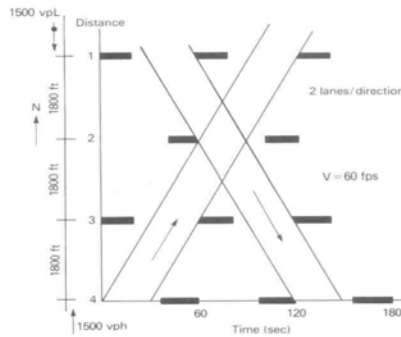
Figure: Bandwidths on a time-space diagram

Bandwidth Concept and Maximum Bandwidth(contd)

- Finding bandwidths
 - It is desired to design one direction with maximum bandwidth subject to some relation b/w bandwidths in two directions
 - Commonly bandwidths in two directions are designed to be in the same ratio as the flows in two directions
 - Offsets may have to move until satisfactory timing plan develops
 - A change in cycle may even be required

Bandwidth Concept and Maximum Bandwidth(contd)

■ Finding bandwidths: A case study



Case study: Four intersections with good progression

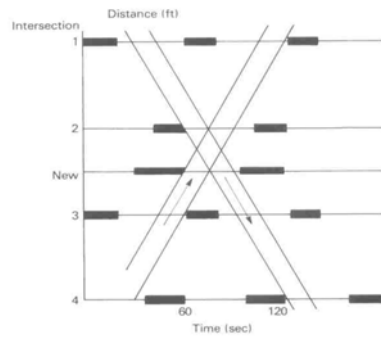


Figure: Effect of inserting a new signal into system

Some Common Progressions

All of these progressions are based on ideal offset equation

$$t_{ideal} = \frac{L}{S} - (Qh + LOSS_1)$$

■ Simple Progression

- All signals are set so that a vehicle released from first intersection arrive at all d/s intersections as they initiate green

■ Forward Progression

- Simple progression is often called forward progression as it results in a green wave that advances with vehicles

Some Common Progressions(contd)

■ Flexible Progression

- Simple progression may be revised two or more times a day, so as to conform to direction of major flow

■ Simultaneous Progression

- All the signals turn green at the same time

Some Common Progressions (contd)

■ Reverse Progression

- Large internal queues may force ideal offset to be negative
- The d/s signal must turn green before u/s for the queue to start moving

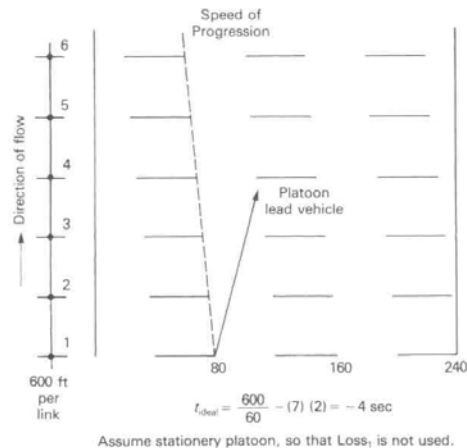


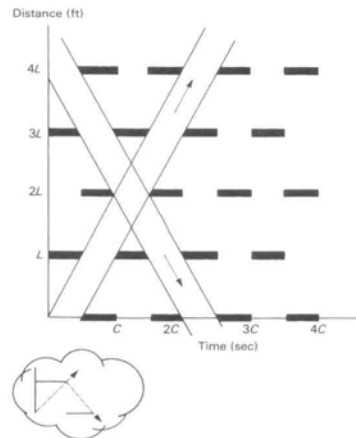
Figure: Illustration of reverse progression

Alternate Progression

- Cycle length is set such that

$$C/2 = L/S$$

For 50:50 splits and Uniform block lengths,
 C = cycle length, sec
 L = block length, ft
 S = platoon speed, fps

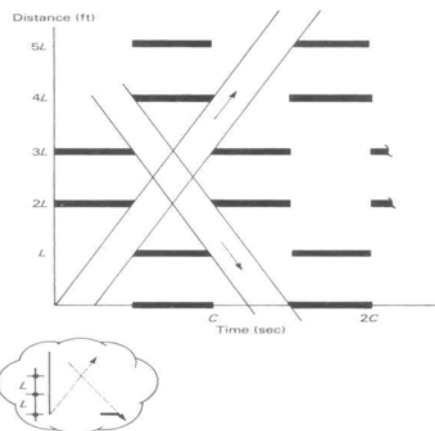


Double Alternate System

- Cycle length is set such that

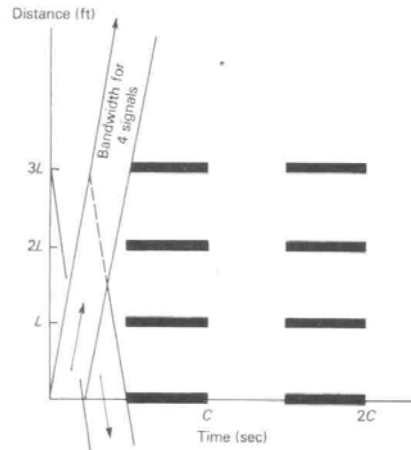
$$C/4 = L/S$$

C = cycle length, sec
 L = block length, ft
 S = platoon speed, fps



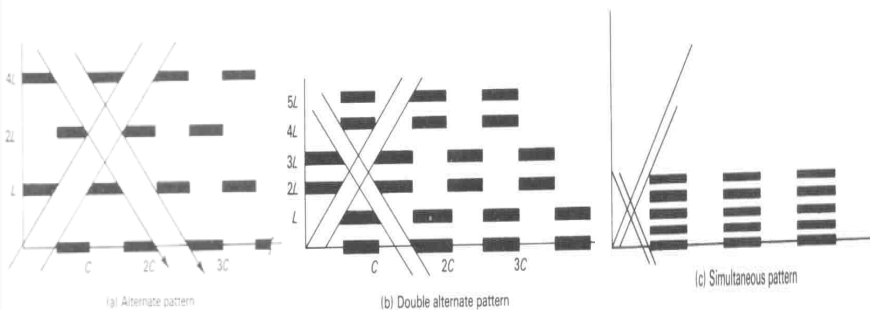
Simultaneous System

- Used for closely spaced signals or high vehicle speeds
- Efficiency depends on number of signals involved
- For N signals, efficiency = $\left[\frac{1}{2} - \frac{(N-1)L}{(S)(C)} \right] 100\%$



Insights of signal Spacing and Cycle Length

- Alternate, double alternate and, simultaneous progressions on the same scale



Insights of signal Spacing and Cycle Length(contd)

- With the arterial changes, progression used may also be changed
- Arterial from a low-density suburban with larger signal spacings, into outlining area of a city, and finally through city's CBDs

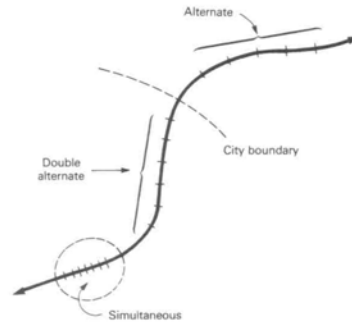


Figure: A hypothetical use of several patterns along the same arterial

Oversaturated Traffic

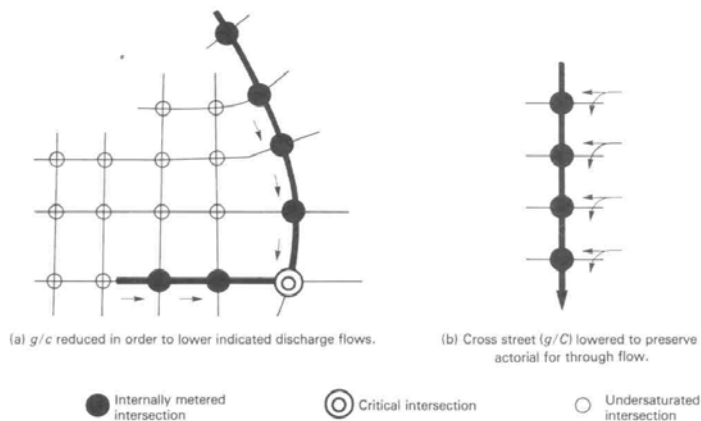
- Extreme congestion is marked by intersection blockage
- Signal based remedies include
 - Metering plans
 - Shorter cycle lengths
 - Equity offsets
 - other

Metering Plans

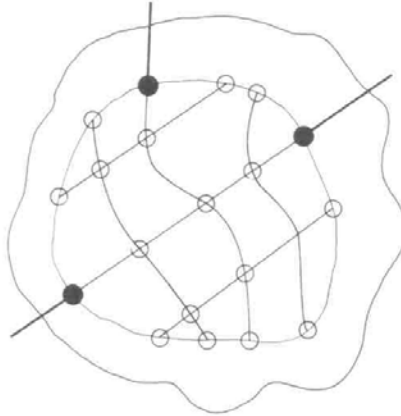
■ Three forms characterized by demand exceeding supply

- Internal Uses control strategies within a congested network to influence distribution of vehicles arriving or departing from a critical location
- External Controls the major access points to the defined systems
- Release Vehicles are stored "offstreet" in such locations as parking garages and lots

Example of Internal Metering



Example of External metering



External Metering: An illustration

Shorter Cycle Lengths

- Cycle length does not substantially increase capacity of the intersection
- Cycle length is given by,

$$C \leq \left(\frac{L}{D}\right)\left(\frac{3600}{V_i}\right)$$

L=available d/s space, ft

D= storage space for each vehicle, ft

V_i =discharge volume per d/s lane, vph

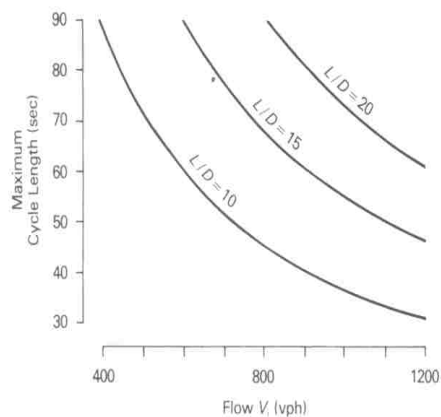


Figure: Maximum cycle length as a function of spacing

Equity Offsets

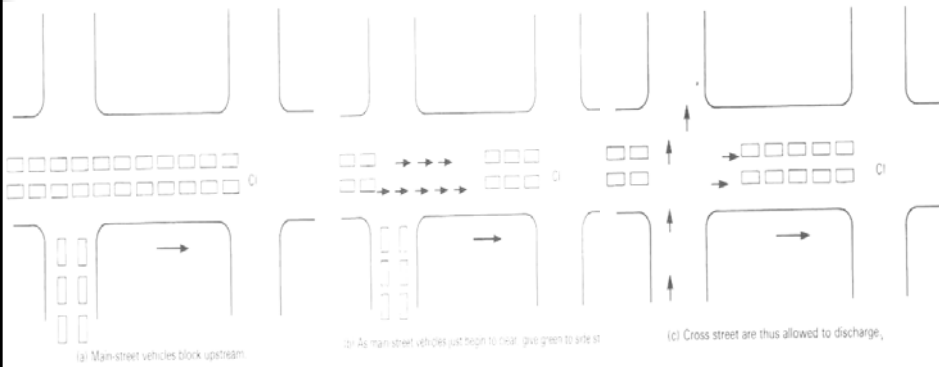


Figure: Concept of equity offsets to clear side streets

Nonsignal Remedies

- Allowing more space
 - Left-turn bays
 - Right-turn bays
- Additional lanes
 - Two-way left turn lanes
 - Tidal lanes

TRANSYT:

A heuristic method for fixed-time control of nonsaturated networks

Robertson, 1969 (substantial extensions & enhancements later)

■ History:

- many applications
- perhaps the most known method
- first test: West London
 - Glasgow: –16% of journey time

■ Main Philosophy

- 1) Model of network (fed with historical demands) → number of stopped vehicles or other (PI Performance Index)
- 2) “Hill-climb” optimization algorithm (heuristic) to minimize the PI
- 3) Decision variables: splits, offsets, cycle time or a combination thereof

TRANSYT (cont.)

■ Initial Signal Settings

- stage structure (given)
- initial choice of split, offset, cycle
- minimum greens

■ Network and Traffic Flow Data

- Network Geometry
- Saturation Flows
- Link travel times
- Turning Rates (constant and known)
- Demands (constant and known)
 - ⇒ time-of-day, steady-state control

TRANSYT (cont.)

The traffic model:

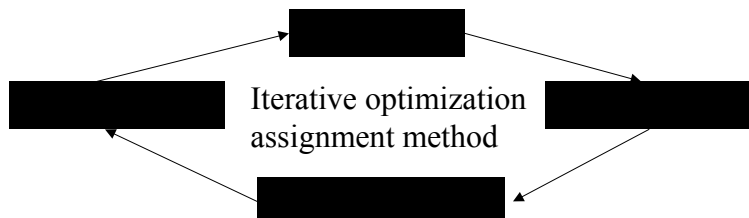
- unique cycle C or $C/2$
- The model consists of:
 - nodes (intersections)
 - links (connecting streets): “platoon dispersion”, i.e. first-order, time-delay system
 - vertical queuing ($\Rightarrow$ no congestion spillback); only undersaturated conditions
 - time discretization $T = C/50$

Problems:

- Dynamic (within-day and day-to-day) demand, turning movements
- Incidents
- Aging
- Oversaturation
- Assignment

Fixed-Time Signal Control and Traffic Assignment

- Signal Control: System Optimum for given demands and turning movements
- Traffic Assignment: User Optimum for given signal control
- (no field implementation available)



SCOOT:

A heuristic method for real-time control of nonsaturated networks

(TRL: Hunt et. al., 1981; later extensions)

- Applied to over 130 cities in U.K. and elsewhere
- Real-time control: avoids aging of fixed-time plans and traffic assignment changes
- Measurements of volumes and occupancies at the upstream end of links
- Central control installation
- "Optimisation" (decision making):
 - splits: few seconds before each stage change, model-based decision (considering each junction+approaches **separately**) on whether the change should be advanced or retarded by a small amount (1-4 s). Criterion: minimise the maximum degree of saturation on the approaches.
 - offset: at every cycle possible change by ± 4 s for each junction separately (based on model-based PIs)
 - cycle: every 5 min according to the most heavily loaded junction in an equal-cycle region

SCOOT (cont.)

■ Limitations:

- Limited traffic responsiveness due to incremental changes (can't do more than 4 sec because the number of combinations would increase significantly)
- No real centralized structure (opposite to claims!).

– Some Recent Developments:

- Bus priority at traffic lights
- Traffic monitoring (ASTRID), e.g. provide historical information in case of detector failure, and other uses.
- Incident detection (INGRID) via simple logic rules

Advanced Real-Time Methods

- **OPAC:** Gartner 1982
- **PRODYN:** Henry and Farges 1989
- **CRONOS:** Boillot, Papageorgiou et al. 1991
- **COP:** Sen and Head 1997
 - No explicit splits, offsets and cycles: specification of switching times over a future time horizon $\Rightarrow J(t_1, t_2, \dots)$
 - Dynamic, realistic models for junctions and links; Discrete problem due to green-red outflows (nonlinear programming algorithms not applicable)
 - Dynamic Programming (PRODYN, COP) Exhaustive Enumeration (OPAC), Global Optimisation (CRONOS)
 - Rolling horizon concept with partly measured (upstream) partly predicted demands; repetition period 2-5 s; horizon > 60 s
 - Basic drawback: exponentially increasing computation time (for DP and EE)
 - Only one intersection (OPAC, PRODYN, COP)
 - Non-global minimum (CRONOS)
- **Reinforcement Learning Approach:** Pringle and Abdulhai, 2000.

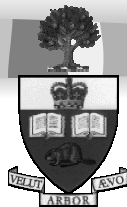
Optimization Based Approaches

- Minimization of total time spent:
 - LP problem
 - tend to give unbalanced queues
- Queue balance:
 - QP problem
- Real time applications:
 - Rolling horizon
- Some Difficulties:
 - Code complexity
 - Real time feasibility

Integrated Network Control

- Main control measures:
 - Ramp metering
 - Signal control
 - VMS
- Drivers perceive the overall network as an entity
- Aims:
 - Harmonise actions of various control measures
 - Avoid antagonistic actions
 - Maximize synergy
- Several de facto implementations but very few methodological approaches:
 - Optimisation-based (LP)
 - Optimal control
 - LQ control: **IN-TUC**

Adaptive Signal Control Using Autonomous Q-Learning Agents



Baher Abdulhai, Rob Pringle,
Grigoris Karakoulas

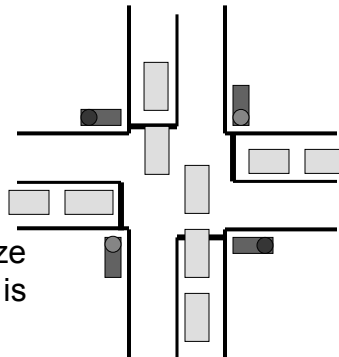
University of Toronto

Outline

- The traffic signal control problem
- Advantages of reinforcement (Q) learning
- Case study – an isolated traffic signal
- Extension to 2-dimensional road networks
- Implementation considerations
- What's next?

The traffic signal control problem: An isolated signal case study

- Need to control time when signal changes indication
- Changes back after fixed cycle length (2-phase signal)
- Objective is to minimize vehicular delay which is proportional to queue lengths



Advantages of reinforcement (Q) learning?

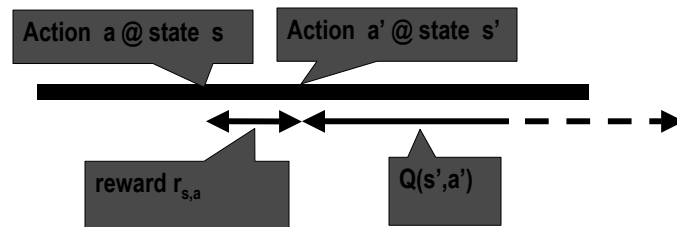
- Learns from ongoing exploration and experience
- Model-free – does not require pre-specified models (eg. for traffic prediction, platoon dispersion)
- Unsupervised - does not require exemplary supervision
- Adaptive to environmental changes
- Optimizes for unique link and network conditions
- Readily accommodates different objective functions
- Permits on-line, real-time enhancement of control policy

The one-step Q-learning algorithm

- Look at it in the context of an isolated traffic signal
- Basic terminology:
 - State – the current state of the environment
(5 elements - queue lengths on intersection approaches plus time elapsed in current phase)
 - Action – the control decision
(change/not change the current indication)
 - Reward – a performance measure
(vehicular delay between decision points – actually a penalty)
 - Policy – a sequence of actions
(series of indication change times)

Q-learning for an isolated traffic signal

$$\delta = \alpha_s (Q_t(s, a) - Q_{t-1}(s, a))$$
$$\delta = \alpha_s (r_{s,a} + \gamma_t \min_{a'} (Q_{t-1}(s', a')) - Q_{t-1}(s, a))$$

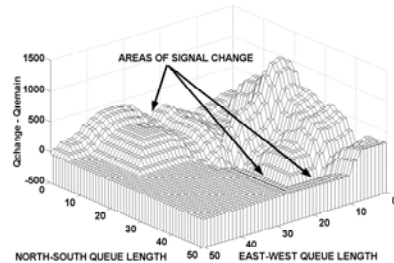


Exploration and training rate

- Adequate exploration required to assure convergence of Q-function in the stochastic traffic environment
- Success with both “softmax” (with annealed temperature function) and ϵ -greedy exploration policies ($\epsilon = 0.9$)
- Training rate (step-size) must be annealed gradually to assure convergence of Q-function
- Rate varied in inverse proportion to number of visits

Function approximation and generalization

- Requires many visits to each state/action pair to converge on stable Q-function
- Need to generalize Q-values through function approximation
- In this case, CMAC's we used:
 - one for decision to change
 - one for decision to remain

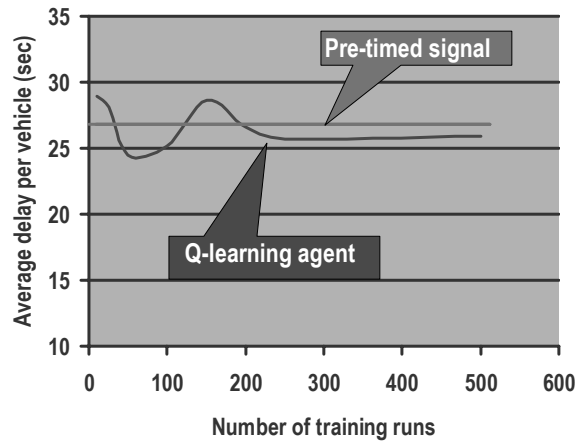


How well does it work?

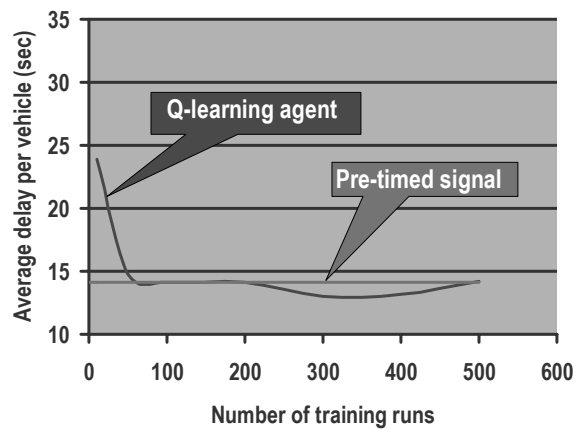
- the isolated signal case

- Tuning the Q-learning agent required an initial series of tests
- The following graphs show the performance of the Q-learning agent, starting from an untrained state under different traffic volume scenarios:
 - Uniform traffic volumes on the intersection approaches
 - Constant-ratio traffic volumes
 - Variable traffic volumes

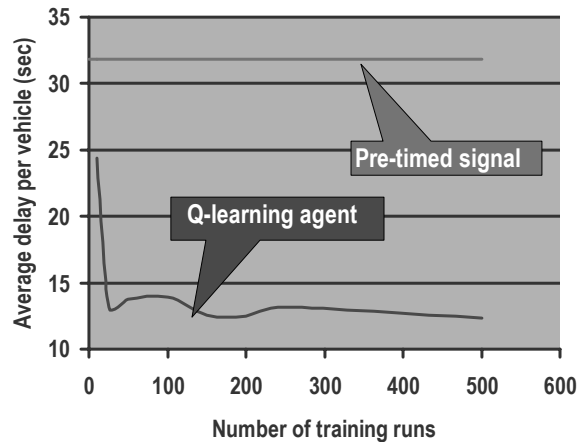
Performance with uniform traffic volumes



Performance with constant-ratio traffic volumes



Performance with variable traffic volumes

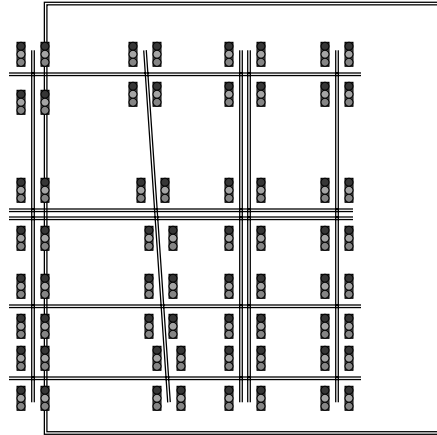


Preliminary conclusions – isolated signal case study

- Under traffic conditions amenable to pre-timed signal control, the Q-learning agent performs comparably
- Under more variable traffic conditions, the Q-learner demonstrates marked superiority
- These tests demonstrate that the reinforcement learning approach works for a simple example – they do NOT show how well it might perform in more complex situations relative to state-of-the-art approaches such as SCOOT

Extension to a 2-dimensional road network

- Need to control:
 - phasing plan
 - phase lengths
 - cycle lengths
 - signal offsets
- Coordination challenges:
 - 2-dimensional network
 - uneven block lengths
 - bi-directional traffic
 - potential for gridlock



Modifications to the problem formulation for the 2-dimensional road network

- Local and communicated “global” state information
- Action becomes the time at which the next change will be made – decision made at time of previous change
- Cycle length not fixed – depends on phase lengths
- Local and communicated “global” rewards – can be weighted differently

Implementation considerations

- In practice, Q-learning agents would be pre-trained on a simulator to ensure reasonable control decisions at the outset
- Once deployed, the Q-learning agents would continue or resume training as necessary to optimize for local conditions and to respond to change (“continuing education / life-long learning”)
- This approach is best implemented with video-based queue sensing using pattern recognition technology
- Need contingency plan to deal with loss of communication

What’s next?

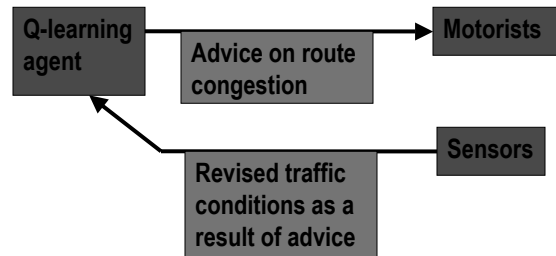
Signal system coordination

- Extension to a network where multiple Q-learning agents would cooperate towards an optimal system control strategy
 - investigation of a linear system currently underway
 - followed by application to a 2-dimensional road network
- Some issues being addressed in the multi-agent framework:
 - definition and weighting of local and global “rewards”
 - extent and nature of communication between agents
 - individual vs collective learning by Q-learning agents
 - on-line updating and review of control decisions

What's next?

Integration with dynamic route guidance

- Once proven in a network situation, integration of adaptive signal control using Q-learning agents with dynamic route guidance will be pursued
- Two-way interaction:



Integrated Multi-jurisdictional Adaptive Corridor Control

- Network is multi-jurisdictional
- Traffic is not
- Integrated control uses:
 - Signals on arterials
 - Ramp metering
 - CMS for re-routing
- Objective:
 - Optimal combined settings for minimum travel times
- Case study:
 - Gardiner/Lakeshore corridor near the Humber under a lane-blocking incident scenario
 - 14% less time
 - 18% better speeds
 - Better numbers for 2-lane blocking incident

