

**Overview:**

Nearly all interconnected traffic routes require a form of traffic control; in moderately congested traffic routes, the traffic light has become ubiquitous and the most effective traffic controller. These signaling systems are incredibly safe and deployed across the U.S. To deploy, designers must determine the length of phases (red/green/yellow) in a light cycle. At the heart of this design problem is determining the optimal time for a phase – exactly which light should be green/red and for how long?

The industry standard for cycle timing is *capacity and critical movement analysis* [1]. In this technique, an intersection's capacity and expected volume are first estimated using measurement data or forecasted with other methods for each critical movement (ex: a Northbound left, a Southbound right, an eastbound straight). The light cycle timings are then allocated according to road geometry, direction of traffic, time of day, and expected traffic flow. Figure one lays out a typical timing diagram for a single approach.

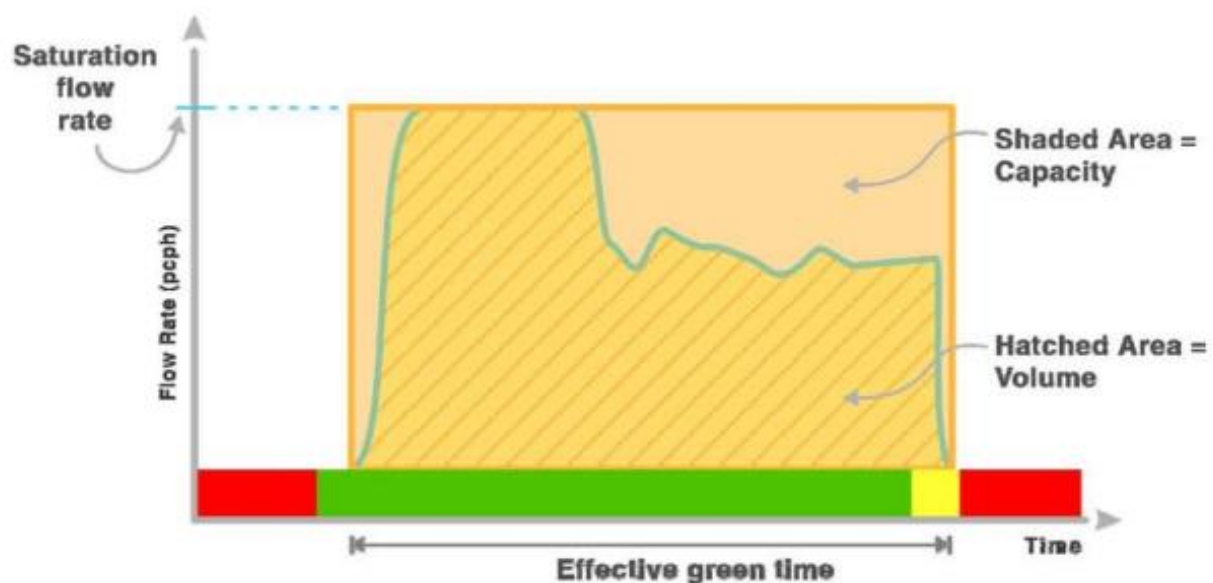


Fig 1. A typical intersection timing diagram for a single approach in a red/green phase. When an approach sees a green signal, the queue clears first (at the saturation flow), then any incoming traffic is let through until the phase time ends. The light then turns yellow and red [1].

This analysis obviously falls short when estimates are incorrect and may produce sub-optimal results. Mismatch between estimate and practice are costly: unnecessary delays add in travel/shipping costs, reduce safety, increase emissions, and consume energy. In some major cities,

timing mismatch costs millions of dollars each year [4]. To combat this inherent flaw, some advanced techniques have been deployed in the past three decades, including phase recalls and improved simulation/time-of-day cycle modelling. In phase recalling, a vehicle is detected, and the current cycle's phase is interrupted to accommodate detected traffic [3]. These detection techniques include induction sensors, pressure sensors, infrared sensors, or camera/computer vision systems [2].

Extending the advancements of vehicle detection techniques, reinforcement learning, coupled with a suitable traffic model, offers a potential for improved traffic flow without relying on estimates. These RL techniques offer an advantage of being able to learn a policy based on the exact state of the system rather than a preconceived estimate. While RL techniques were first applied to this problem in 2011, the problem is finally getting more traction and seeing a variety of approaches taken [5] [6]. This rich intersection between reinforcement learning and classical traffic design is well-poised for a final project and can demonstrate some of the techniques learned in class.

### **Project Outcomes:**

In this project, the following (rather simple) traffic model will be assumed:

- Traffic flows N/S and E/W at an intersection. The volume will have a stochastic component and differ between the directions.
- A two-phase system will be designed and no turns will be permitted.
- The number of cars in the queue at each approach can be reliably counted.
- There will be a multi-second yellow light phase between phases.
- The vehicles will be a unit length,  $L$ , and travel at a constant velocity unless stopped.

Baseline goal/main approach:

- Construct an MDP of the traffic model for a two-phase cycle. A (very) rough MDP construction is below:
  - States: Occupied/unoccupied spots in a traffic queue
  - Actions: N/S green light, E/W green light, yellow transition light (red on non-green directions)
  - Reward: Negative reward correlating to the length of the traffic queue
  - Transition: Allow a queue to clear one vehicle in a time step during green lights and stochastically accumulate vehicles as they enter the queue at a stopped light
  - Discount reward: TBD.
- Using a deep-RL technique (Q-learning or Maximum-likelihood RL) to construct a policy that minimizes the traffic queue or traffic wait time. This will be implemented directly in Julia by adopting homework assignments.

Medium goal:

- Compare the RL-policy to state-of-the-art traffic planning techniques for a comparable system and volume.

- Compare different exploration techniques/tuning the model to gain insight in how the RL is constructing the policy.
- Flow visualization.

Stretch goal:

- Extend the state/action space to two *sequential* intersections.
- Extend the state/action space to four intersections, arranged like a grid to simulate a city-like environment.
- Modify the MDP to consider coordinated cycle timing (ex: found typically in downtowns, where all the lights on an arterial turn green at once and side streets all remain red).
- Visualization of the four-intersection model.

### Citations:

- [1] "Traffic Signal Timing Manual," US Dept of Transportation, Federal Highway Administration, Washington, DC, 2008.
- [2] "Traffic Detector Handbook," US Dept of Transportation, Federal Highway Administration, Washington, DC. , Oct, 2006.
- [3] "Traffic Signal Design - Detection. TDOT Traffic Design Manual," TDOT, Nashville, TN., 2016.
- [4] B. T.-S.-H. Associates, "The Benefits of Retiming/Rephasing Traffic Signals in the Back Bay," Boston Transportation Dept, Boston, MA, 2010.
- [5] L. B. Y. T. Chin, "Exploring Q-Learning Optimization in Traffic Timing," in *Third International Conference on Computational Intelligence, Communication Systems and Networks*, 2011.
- [6] W. L. Han, "Leveraging reinforcement learning for dynamic traffic control: A survey and challenges for field implementation," *Communications in Transportation Research*, vol. 3, 2023.