Reward Functions for Real-World Pedestrian and Vehicular Intersection Control through Reinforcement Learning

Alvaro Cabrejas-Egea & Colm Connaughton

WARWICK
THE UNIVERSITY OF WARWICK

centre for complexity science

EPSRC
Engineering and Physical Sciences Research Council

VIVACITY

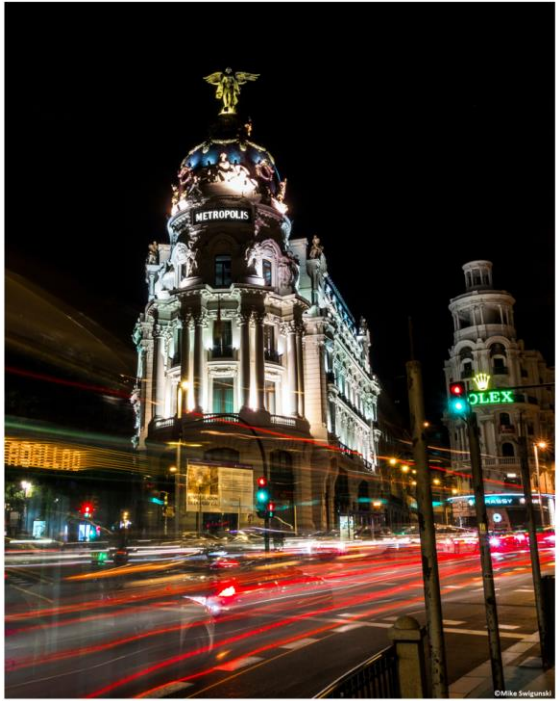
Times of Change in Urban Traffic Control

Traditional TSCs

- Fixed Time Controllers
- Vehicle Actuated
  - System D
- Adaptive Controllers
  - MOVA
  - SCOOT
- Area Controllers
  - Surtrac

New Generation of TSCs

- Extra Computing Power
  - GPU-based computation
  - Image analysis
- New Generation of Sensors
  - Vision Based
- New available measurements:
  - Vehicles in Queue
  - Individual Delays
  - Individual Speeds
  - Individual Positions
  - Class of vehicle



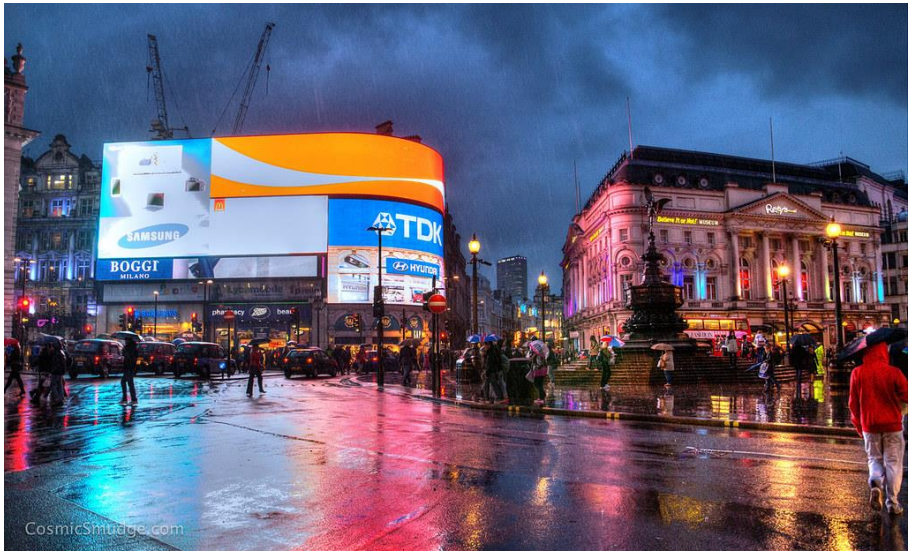
A tale of pedestrians

Plenty of literature on RL for UTC

- Focused on Vehicular Intersections

Specific Requirements for Pedestrians

- Impact on Learning
- Multi-modal optimisation
- Need for more sophisticated data



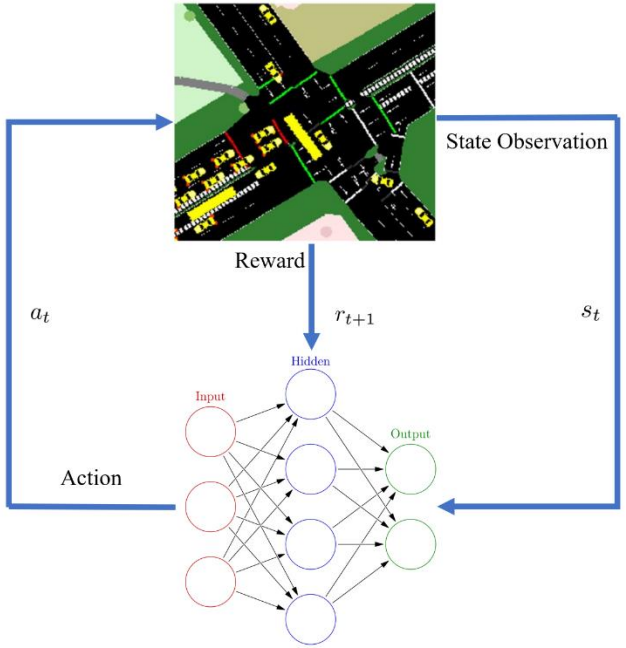
Sensing in the Real World

van
van
van
car
car
pedestrian
pedestrian
pedestrian
pedestrian
pedestrian
BUS
LANE

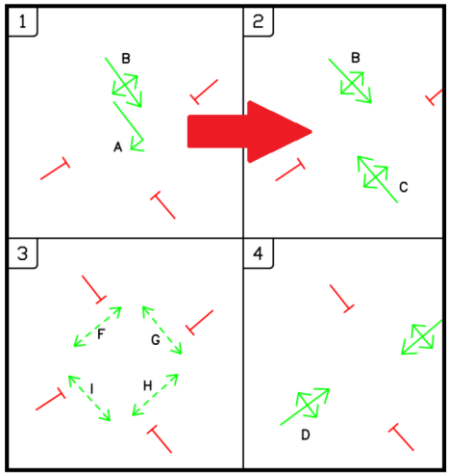
Reinforcement Learning for UTC

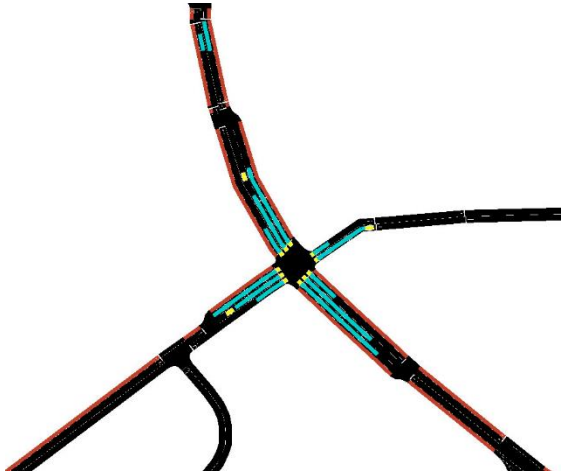
Components:

- State estimation (sensing)
- Choosing an action (decision)
- Reward Calculation (feedback)



Deployment Junction





Reward Functions

Quantities

- Queues
- Waiting Time
- Delay
- Average Speed
- Throughput

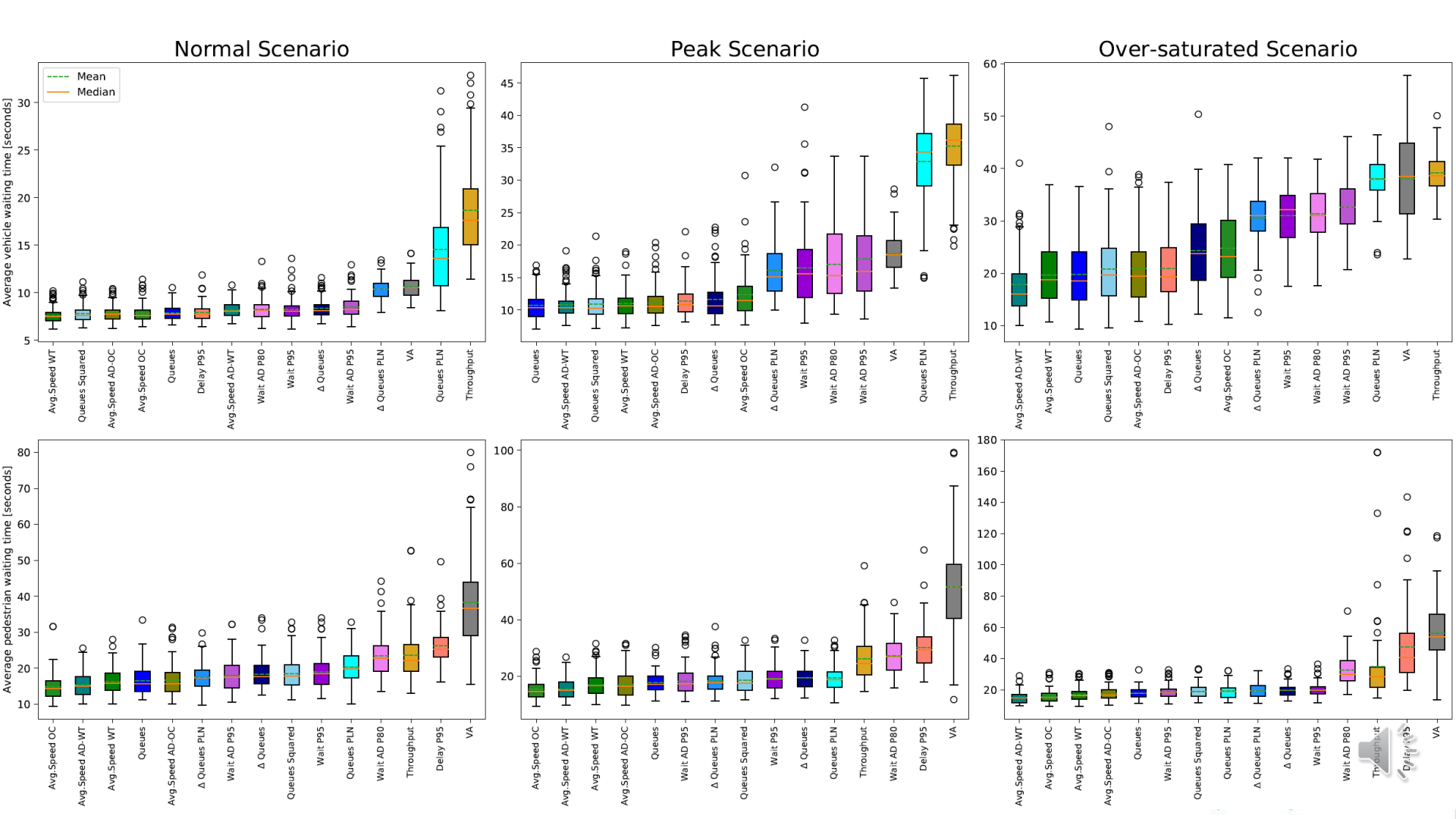
Transformation Components

- Instant measurement
- Variation (Δ)
- Demand estimation
- Phase-Length Normalisation
- Squared Variables

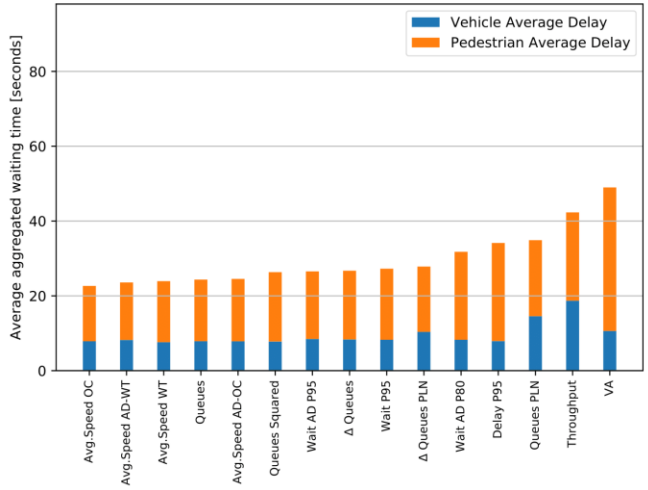
Reward = a * Reward Vehicles + b * Reward Pedestrians

Agents, Training and Testing

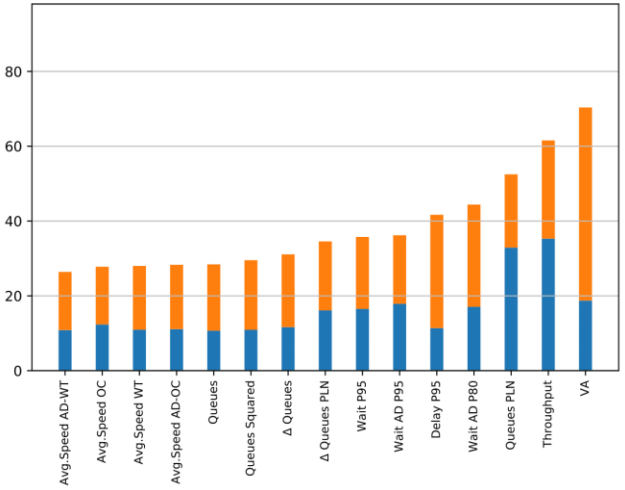
- Standard DQN implementation
- State = Occupancy, 12 second buffer at δ=0.6s
- 2 Hidden layers (sizes 500, 1000), using ReLU.
- Optimized with ADAM, $\alpha = 10^{-5}$; $\gamma = 0.8$
- Trained 1500 episodes of 30 minutes
- Testing 100 copies of 3 demand levels


Normal Scenario
Peak Scenario
Over-saturated Scenario
Mean
Median
Average vehicle waiting time [seconds]
Average pedestrian waiting time [seconds]
Avg.Speed WT
Queues Squared
Avg.Speed AD-OC
Avg.Speed OC
Queues
Delay P95
Avg.Speed AD-WT
Wait AD P80
Wait P95
Δ Queues
Wait AD P95
Δ Queues PLN
VA
Queues PLN
Throughput

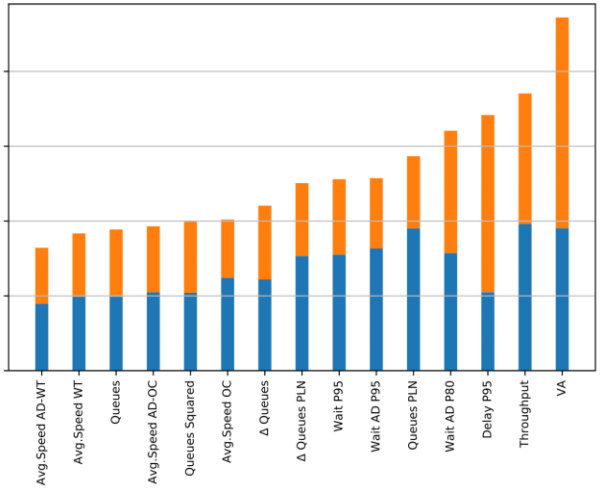
Normal Scenario



Peak Scenario



Over-saturated Scenario



Future Work

- Extension to multiple intersections
- Granular per-entity measurements
- Multi-modal control

Thank you!

a.cabrejas-egea@warwick.ac.uk