Reinforcement Learning
Architecture 10/09
ORCS 4529: Fall 2025

Aksel Kretsinger-Walters, adk2164
adk2164@columbia.edu
Due: 10/09/2025 11:59 PM ET

Modeling Setup

State

Discrete-Time Markov Decision Process (MDP)

Time steps are 1 minute long

Environment

- Static class - no state changes
- Embedded into l_2
- Graph -
  - Nodes - intersections
  - Edges - roads
- Vertices -
  - Static class - no state changes
  - Embedded into l_2

Fleet

- n=10? vehicles
- Per vehicle:
  - Vehicle ID
  - Location
  - Battery level (percentage)
  - Status (available, en route to pickup, with passenger, en route to charging station)

Demand I'm a little unsure about whether this should be *explicitly* modeled in the state or "learned" by the agent. I'm learning towards the former for a few reasons

(a) It's relatively independent of the actions of the agent

(b) It can be estimated from historical data

(c) Both assumptions above are likely to be true in the real world

(d) It greatly simplifies the learning problem

If we do explicitly model it

- Location
- Expected demand (number of requests) in next time step

Ride Requests

- $n_v \sim \text{Poisson}(\Lambda)$ new requests per time step (this lives in Demand class)
- Each request:
  - Request ID
  - Pickup location
  - Dropoff location
  - Request time
  - Status (pending, assigned, completed, cancelled)

Active Rides I'm a little less sure about this one, but I think that this class should mostly serve as an enriched tuple between a vehicle and a ride request

- Matched Vehicle ID
- Request ID
- Status (en route to pickup, with passenger)
- Time to pickup (minutes)
- Duration of ride (minutes)
- Estimated remaining time (minutes)