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

- Locations will be represented either/or as latitude/longitude coordinates or as nodes on a graph (any point on vertex is proportional distance between endpoints)

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

-

Ride Requests

- $n_v \sim \text{Poisson}(\Lambda)$ new requests per time step

- 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)

We model the state as a tuple (t, F, R, D) where:

- t is the current time step (discretized to 10 minute intervals)

- F is a list of all vehicles in the fleet, where each vehicle is represented as a tuple $(id, loc, status, battery, time\_to\_maintenance)$

    - id is a unique identifier for the vehicle
    - loc is the current location of the vehicle (latitude, longitude)
    - $status$ is the current status of the vehicle (available, en route to pickup, with passenger, en route to charging station, en route to maintenance)
    - $battery$ is the current battery level of the vehicle (percentage)
    - $time\_to\_maintenance$ is the time remaining until the vehicle requires maintenance (in hours)

- R is a list of all ride requests, where each request is represented as a tuple $(id, pickup\_loc, dropoff\_loc, \imath$

    - id is a unique identifier for the request
    - $pickup\_loc$ is the location of the pickup (latitude, longitude)
    - $dropoff\_loc$ is the location of the dropoff (latitude, longitude)
    - $request\_time$ is the time the request was made
    - $status$ is the current status of the request (pending, assigned, completed, cancelled)

- D is a list of all demand hotspots, where each hotspot is represented as a tuple $(loc, expected\_demand)$

    - loc is the location of the hotspot (latitude, longitude)
    - $expected\_demand$ is the expected number of ride requests in the next time step