ParkRL: Learning Generalizable Parking via DRL

Raine Cui • Georgia Lin • James Liu • Tianyao Ren

CMPT 729 Reinforcement Learning

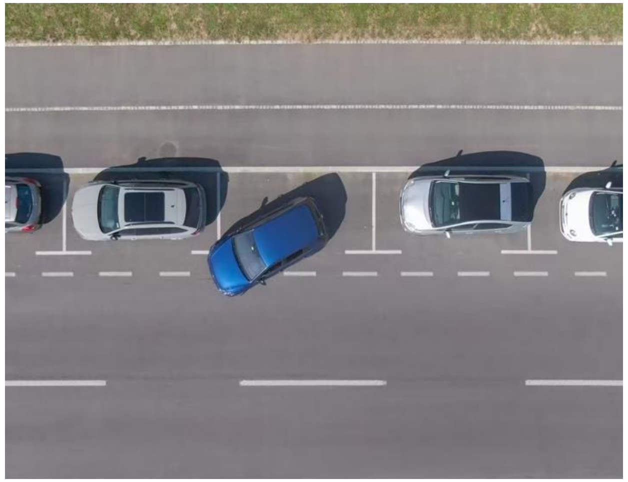
December 1, 2025

Introduction

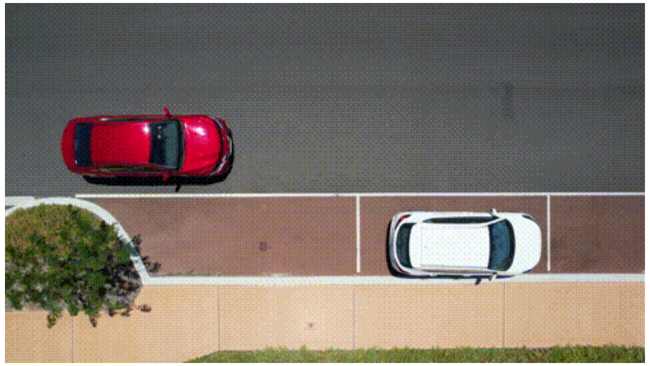
Perpendicular Parking (90°)

Angled Parking (45°)

Parallel Parking (0°)

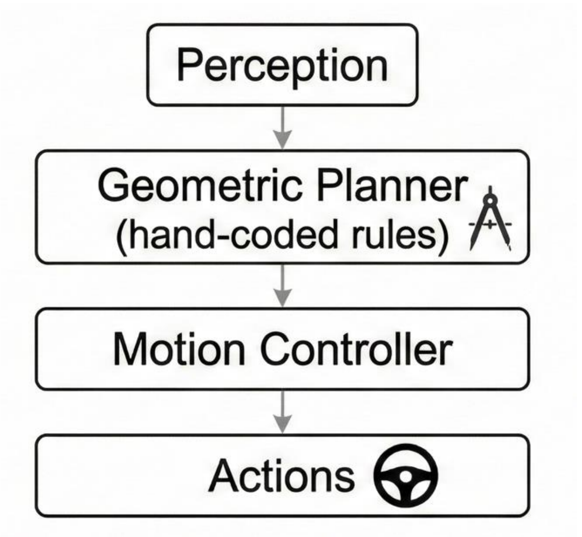


- Challenges
  - **Non-holonomic**: Can’t slide sideways
  - **Continuous control**: Throttle + steering
  - **Dense obstacles**: Cars, walls, barriers

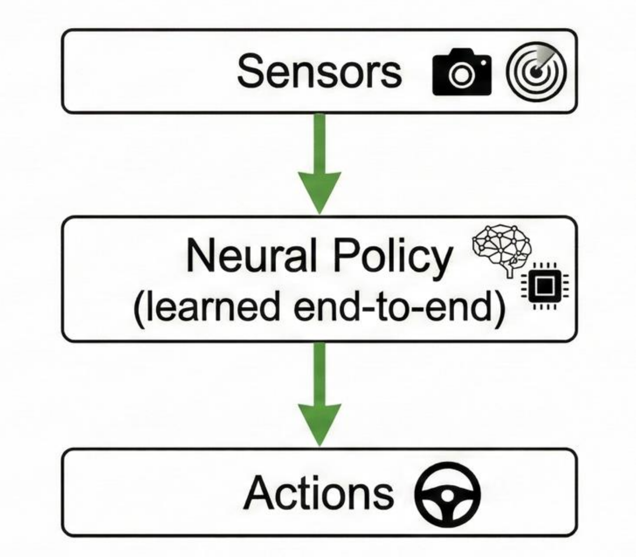


Related Work

- **Traditional Methods**
  - × Scenario-specific tuning
  - × Slow replanning
  - × **Poor generalization** to new layouts



- RL Approach
  - ✓ One policy for ALL scenarios
  - ✓ Fast O(1) inference time
  - ✓ Generalization via randomized training



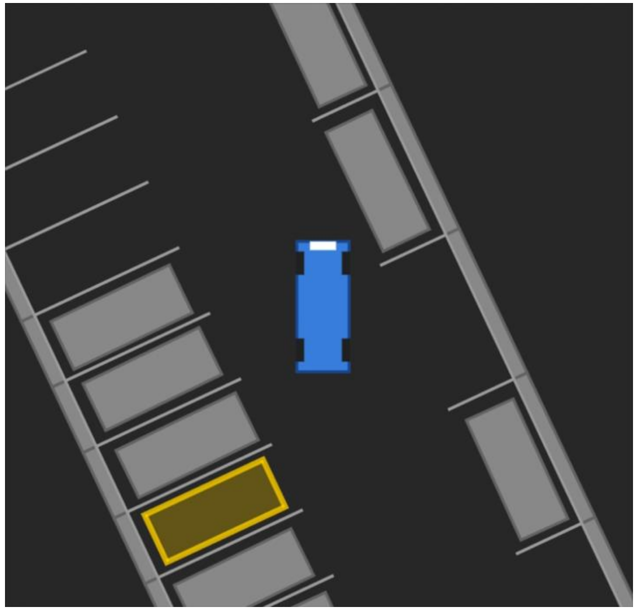
Method Overview

- Random Parking Lot → Observations → Neural Policy → Actions

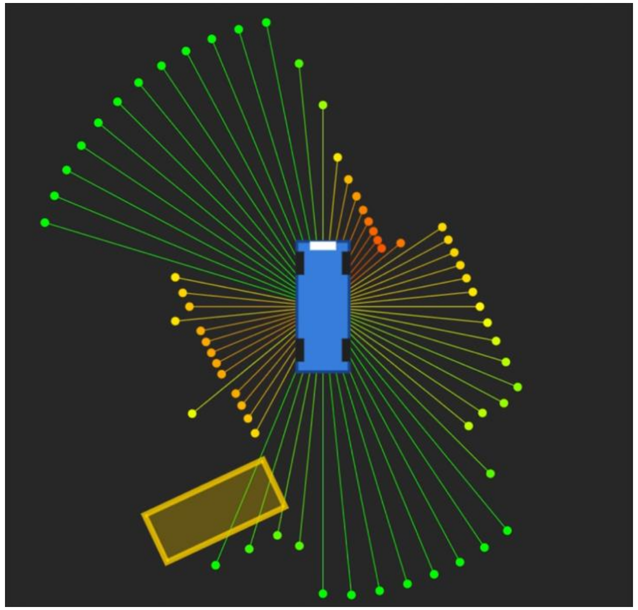
- **Observations**
  - 64-ray LiDAR
  - Self speed & steering
  - Target pose

    Ego-centric coordinates

    ➔ Generalization

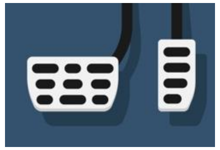


Parking Lot Scene

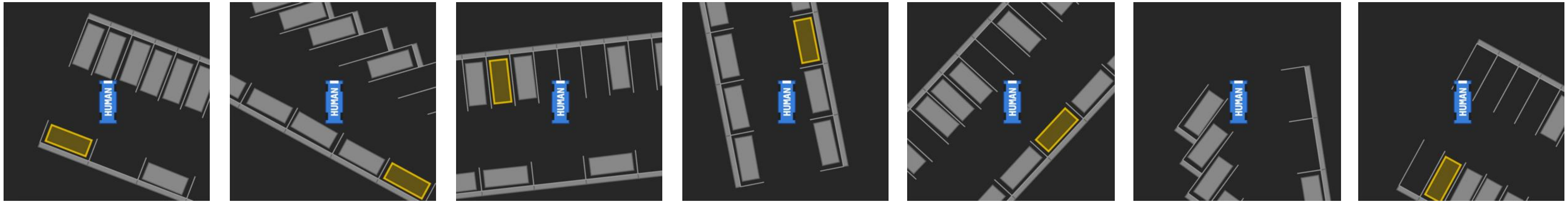


Observations

- **Actions**
  - Target speed
  - Target steering

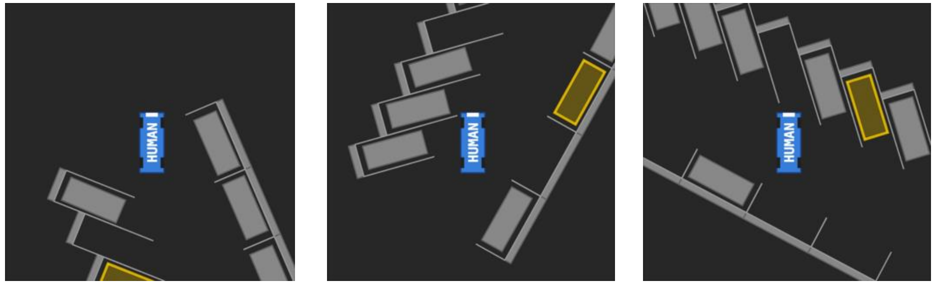


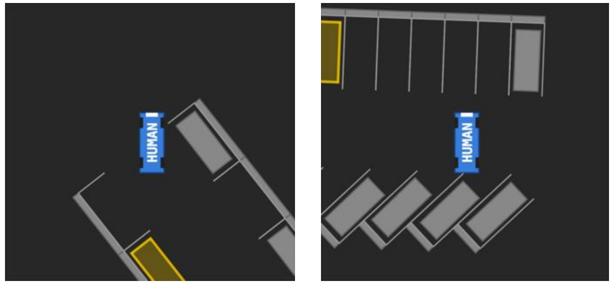
Randomized Training Environments



- **Procedural Generation**
  - Random parking types (90°/45°/0°)
  - Probabilistic obstacles (50% density)
  - Random barriers (walls/curbs)
  - Variable road width (6-10m)

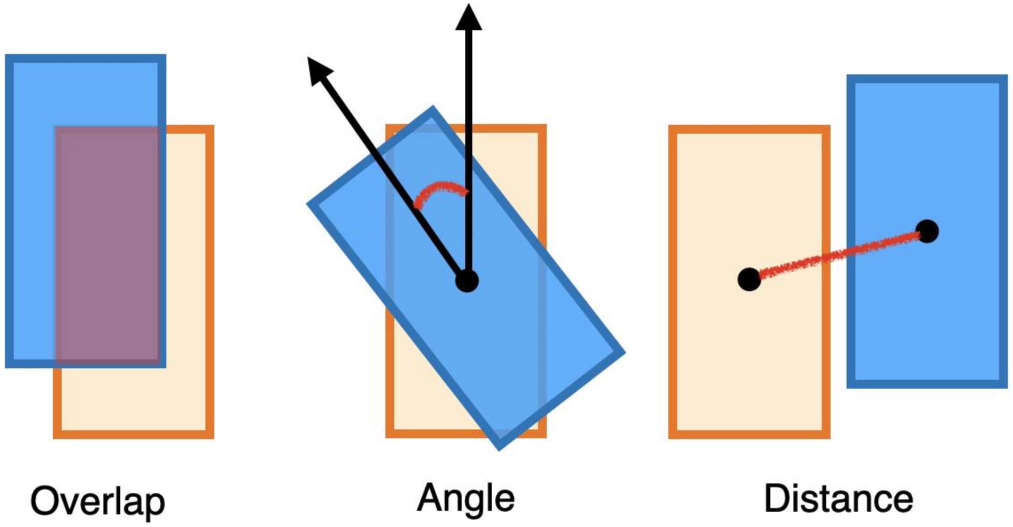
- **Every episode = different parking lot**
  - Policy must generalize, can't memorize scene!





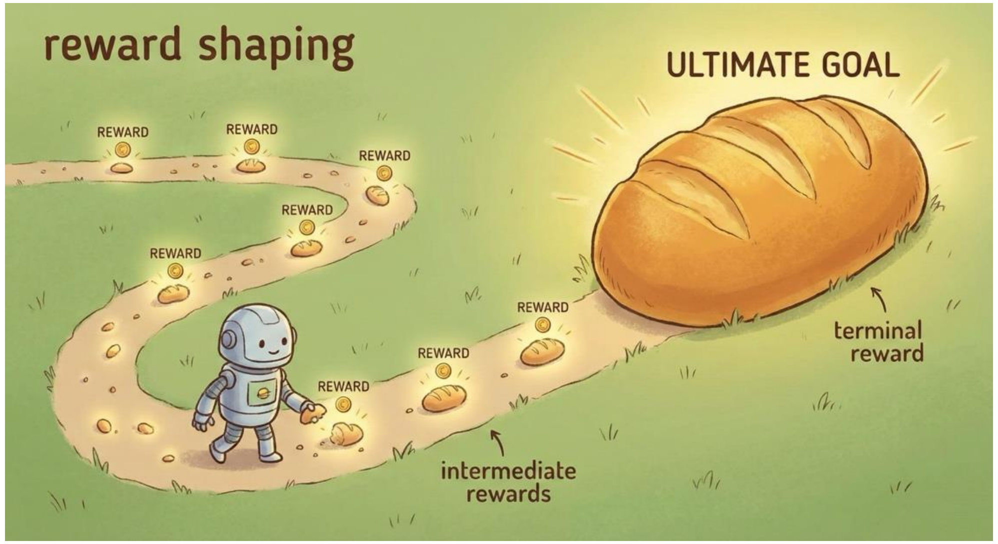
Reward Design

- **Successful Parking**
  - Up to +100 pts
  - Harmonic mean of (Overlap, Distance, Angle)



- **Collision Penalty**
  - $-(1 + v^2)$ pts

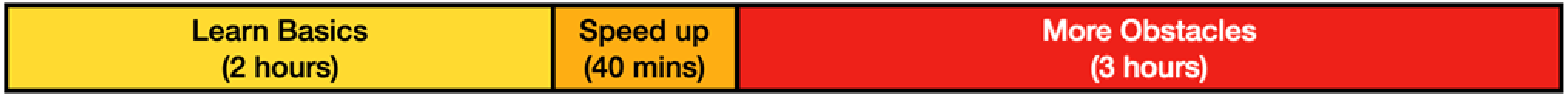
- **Reward Shaping**
  - **Distance**: +1 pt per meter closer
  - **Time**: -0.5 pts per second
  - **Gear-Shifting**: -1 pt each time



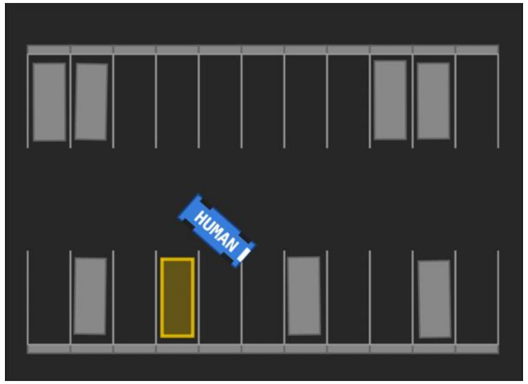
breadcrumbs leading to final bread

Curriculum Learning Strategy

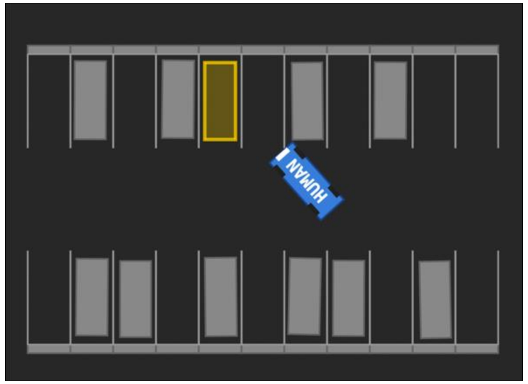
- Three-Stages



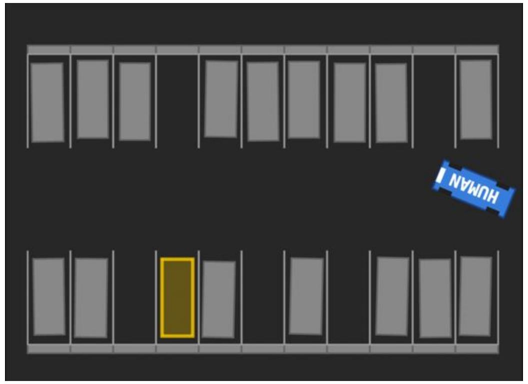
1. **Learn basics (2h)** (low speed, few obstacles)
2. **Increase speed (40min)** (increased speed 1 m/s → 3 m/s → 5 m/s)
3. **More obstacles (3h)** (density 20% → 40% → 80%)



obstacle density 20%



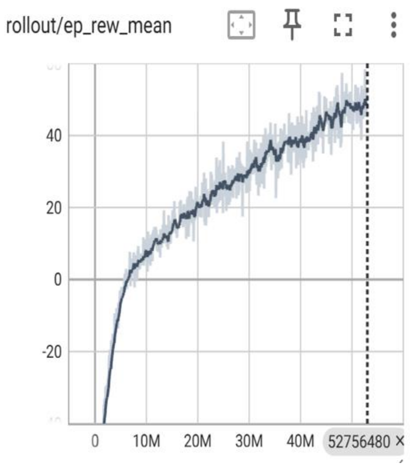
obstacle density 40%



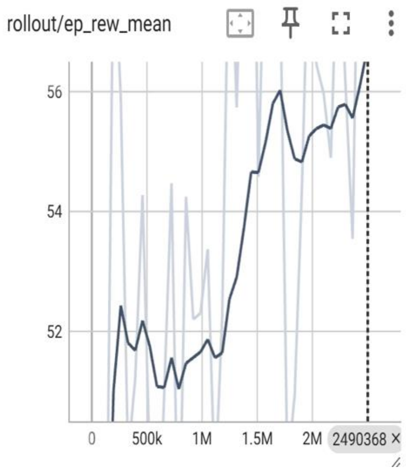
obstacle density 80%

- **Necessity**: Will be demonstrated in ablations

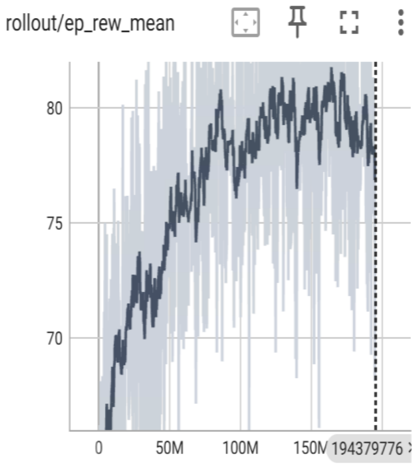
Training Results



Stage 1
-50 → 45

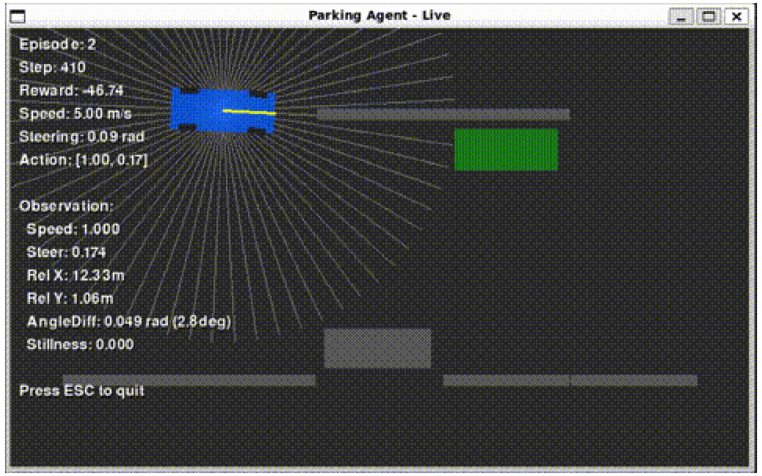


Stage 2
45 → 57

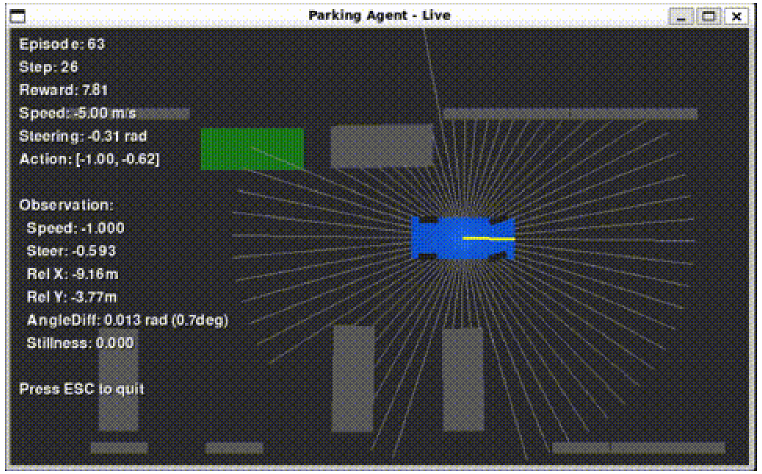


Stage 3
57 → 80

- Converged Performance
  - Avg. Reward: ≈ **80 pts**
  - Success rate: ≈ **96 %**
  - Parking speed: ≈ **3.7 m/s**
  - Collision every 7 episodes



Early-stage model struggles to park

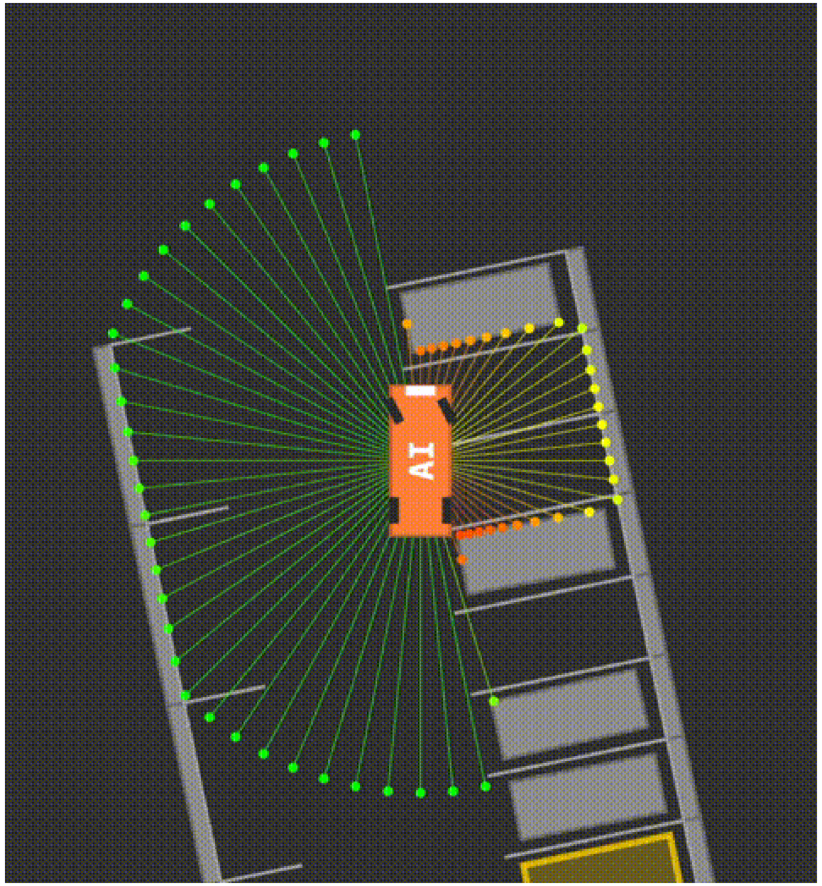


Converged model parks quickly

Interactive Demo

- https://bit.ly/729demo (or https://pkucuipy.github.io/rl-final)

- **Interactive Features**
  - **J/K/L** – Switch LiDAR Render Modes
  - **T/Y** – Switch first-person camera
  - **M/N** – Disable/Enable AI

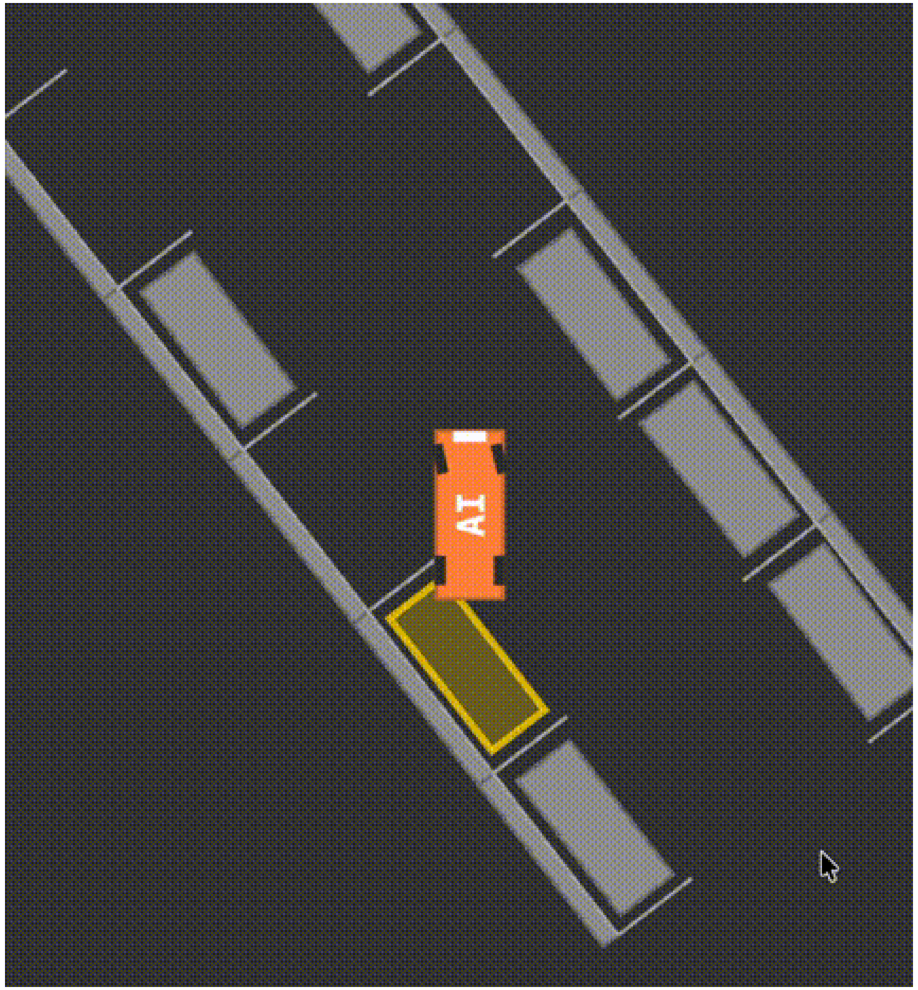


Interactive Demo

- https://bit.ly/729demo (or https://pkucuipy.github.io/rl-final)

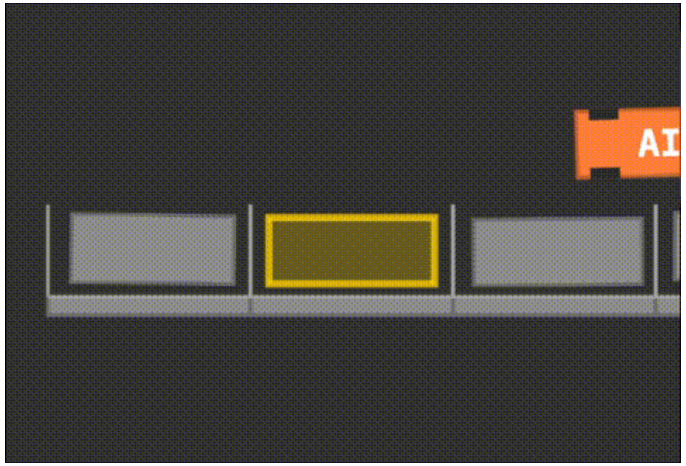
- **Interactive Features**
  - **J/K/L** – Switch LiDAR Render Modes
  - **T/Y** – Switch first-person camera
  - **M/N** – Disable/Enable AI

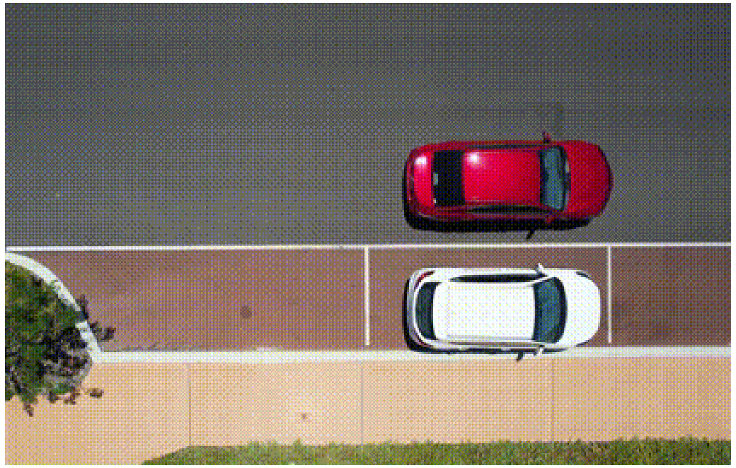
- **AI-Copilot Mode!**



Emergent Behaviors

- What Did the Agent Learn?

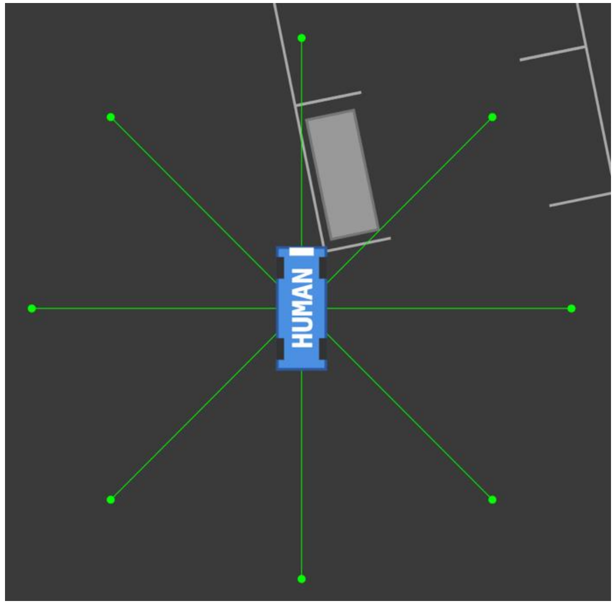


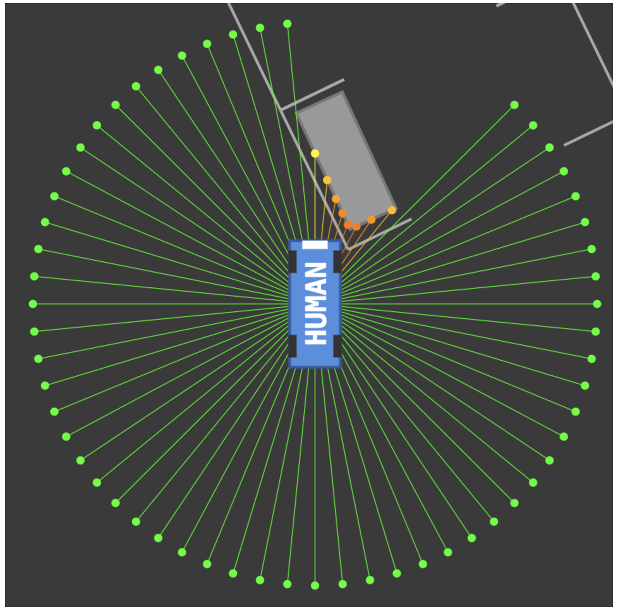


- Discovered Strategies
  - ✓ Reverse bay parking
  - ✓ Reverse parallel parking
  - ✓ Matches human driving school techniques!

Ablation 1 – LiDAR Resolution

- Why 64 rays? Can we use Fewer?
- 8 Rays Failure Case: Blind spots!

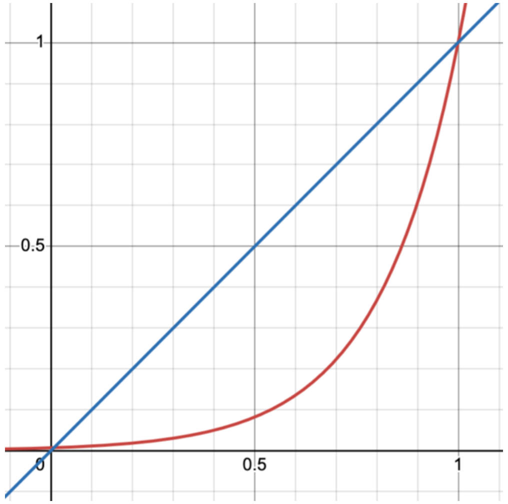




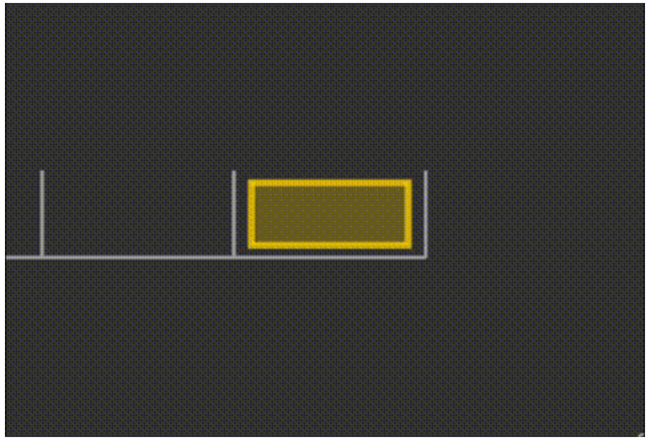
- Problem: Without memory, blind spots cause:
  - Unexpected collisions
  - Unstable, jumpy control policies

Ablation 2 – Reward Exponential Decay

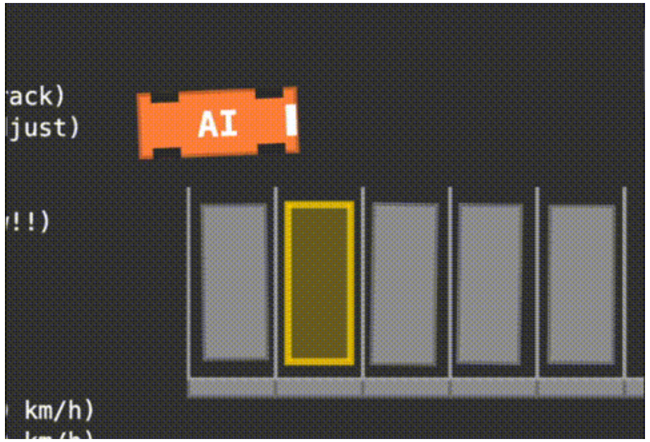
- No Decay ***vs*** Exponential Decay
  - No decay: “Good enough” at 45° → mediocre parking
  - Exp(-5x): Precision required → perfect alignment



- Harder rewards → better strategies
  - Agent discovers reversing maneuvers!
  - Counter-intuitive: Higher converged reward ≈ 50 → ≈ 80!



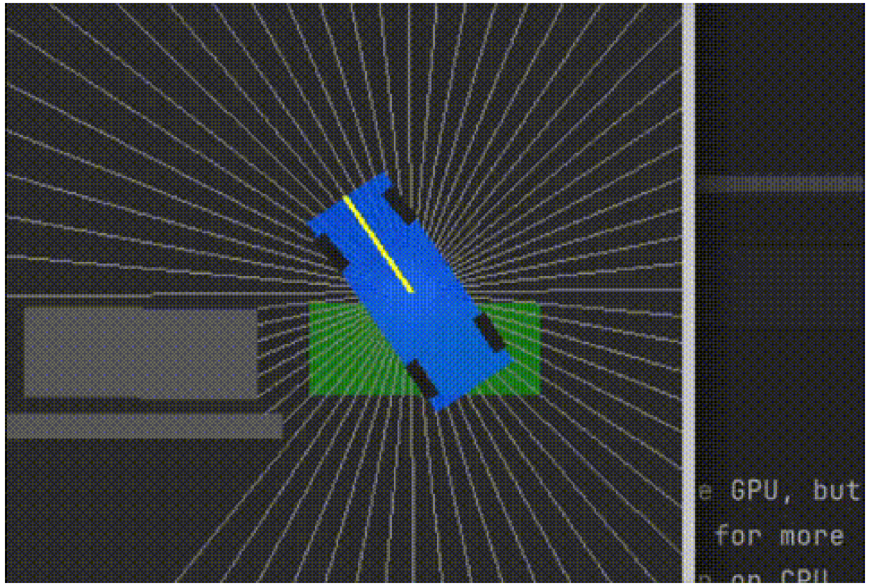
mediocre parking



reversing maneuvers

Ablation 3 – Is Curriculum Learning Necessary?

- From-scratch: doesn’t converge
  - Never learns to stop
  - Only learns obstacle avoidance



Never learns to stop

- With Curriculum: converges in 6 hours
  - Successfully learns all behaviors

- Conclusion: Curriculum essential for multi-objective tasks

Limitations & Future Work

- **Soft Collision Constraints**
  - Current: Penalty-based (no hard guarantee)
  - **Idea**: Beam search over sampled rollouts
- **No Temporal Memory**
  - Current: Stateless policy (may re-explore)
  - **Solution**: Recurrent architecture (LSTM/Transformer)
- **Limited Model Capacity**
  - Current: Small 2-layer MLP
  - **Direction**: Scale up, use pre-trained vision features

Demo: https://bit.ly/729demo
or https://pkucuipy.github.io/rl-final

Summary

- ✓ End-to-end RL for generalized parking
- ✓ Novel reward design
- ✓ Systematic curriculum strategy
- ✓ Emergent human-like behaviors
- ✓ Deployment-ready (Python → ONNX → WebJS)

Presented by: Raine Cui • Georgia Lin • James Liu • Tianyao Ren

Thank You!