# CMPUT 655
# Assignment 1

September 5, 2024

Dikshant
1877098

---

1. *Formalize a Roomba Robot cleaning an apartment as an MDP*

**Possible solution:**

1. **States (S)**:
    - Current position of the robot
    - Battery level
    - Room grid cleanliness status
2. **Actions (A)**:
    - Throttle
    - Braking
    - Steering
    - Charging
3. **Rewards (R)**:
    - Cleaning the trash: +1
    - Hitting an obstacle: -10
    - Running out of battery: -10

*Assumption: The room is divided into grids and each grid is having a cleanliness status: 0 or 1, representing whether the grid is cleaned or not.*

**State Representation**:
- *Current Position of the Robot*: The position on a grid or map of the apartment represents where the robot is located. This allows the robot to plan its path,

avoid obstacles, and cover all areas efficiently.
- *Battery Level*: Knowing the battery level is essential for the robot to decide when to return to a charging station. This ensures that the robot doesn't run out of power mid-task, which would be inefficient and counterproductive.
- *Room Status*: This tracks the cleanliness status of each grid cell within the apartment. The robot needs this information to determine which areas still require cleaning and to ensure all cells are cleaned before the task is considered complete.

**Action Dynamics**:
- *Throttle*: This action controls the robot's movement speed. Proper throttle control allows the robot to move quickly when there are no obstacles or to slow down in more cluttered areas.
- *Braking*: Braking is crucial for preventing collisions and ensuring the robot stops when it detects an obstacle or when precise positioning is needed.
- *Steering*: Steering actions control the robot's direction, allowing it to navigate around furniture, reach different areas, and avoid obstacles.
- *Charging*: Charging action directs the robot to return to a charging station when its battery is low. This ensures it maintains enough power to complete the cleaning task without interruption.

**Reward Structure**:
- *+1 Reward for Cleaning the Trash*: A positive reward for cleaning each grid cell or area encourages the robot to focus on its primary task. This incentivizes the robot to seek out and clean dirty areas actively.
- *-10 Reward for Hitting an Obstacle*: A significant penalty for collisions with obstacles (e.g., furniture, walls) teaches the robot to avoid such situations. This discourages careless movements and promotes careful navigation.
- *-10 Reward for Running Out of Battery*: A penalty for running out of battery ensures the robot learns to manage its power levels effectively. The robot is encouraged to return to the charging station before its battery is depleted, ensuring continuous operation and efficiency.

Overall, this reward structure helps the robot learn to balance cleaning efficiently with safety (avoiding obstacles) and resource management (maintaining sufficient battery life).

2. *Formalize autonomous driving as an MDP*

**Possible solution:**

1. **States (S)**:
   - Vehicle data:
     - Position
     - Velocity
     - Acceleration
     - Orientation
   - Other vehicles data []
   - Road data:
     - Speed limit
     - Lane coordinates
     - Curvature
2. **Actions (A)**:
   - Throttle
   - Braking
   - Steering
3. **Rewards (R)**:
   - Drive in a straight line: +3
   - Course progress: +3
   - Collision with traffic participants: -100
   - Jerk: -10
   - Overtake: +1

**State Representation**: The defined states capture all necessary information about the vehicle's surroundings and current status to enable it to navigate the course and complete all laps effectively. The state includes the vehicle's position, velocity, orientation, lane information, and the relative positions and speeds of nearby vehicles and obstacles.

**Action Dynamics**:
- *Throttle*: This action is responsible for accelerating the vehicle in a controlled manner.
- *Braking*: The braking action is designed to decelerate the vehicle, helping it to avoid collisions and adhere to traffic rules or sudden stops.
- *Steering*: Steering actions change the vehicle's direction to align with the desired trajectory. This action ensures the vehicle can navigate curves, avoid obstacles, and stay within lane boundaries.

**Reward Structure**:
- *+3 Reward for Staying in the Lane*: A positive reward for driving in a straight line incentivizes the vehicle to maintain its lane. This reward encourages the agent to learn lane-keeping behavior and avoid drifting or veering off course.
- *+3 Reward for Course Progress*: By rewarding progress towards completing laps, the vehicle is motivated to keep moving toward its goal. This helps the vehicle to follow the designated path and discourages off-course driving.
- *-100 Penalty for Collisions*: A substantial negative reward for collisions emphasizes the importance of safety. By heavily penalizing crashes with other vehicles, pedestrians, or obstacles, the vehicle learns to prioritize avoiding accidents, which is crucial for real-world driving.
- *-10 Penalty for Jerky Movements*: To ensure a comfortable and safe driving experience, a penalty is given for rapid acceleration, sudden braking, or sharp turns. This discourages aggressive maneuvers and promotes smooth, controlled driving behavior.
- *+1 Reward for Safe Overtaking*: A small positive reward for overtaking another vehicle safely encourages efficient driving without compromising safety. This reward helps the vehicle learn to balance the need for speed and progress with the requirement to avoid collisions.

Overall, this reward structure encourages the autonomous vehicle to drive safely, stay in its lane, make steady progress towards completing laps, avoid collisions, drive smoothly, and take advantage of overtaking opportunities when they arise.