MDP: Student Housing

# Overview

This Markov Decision Process (MDP) models designs a scenario revolving around the management of food and interactions with a neighbor in Student Housing. This model captures the constarints and possible actions of decision-making in everyday scenarios.

# States

The MDP contains three distinct states reflecting different scenarios a person might face in the context of food management:

1. **You Have Food**: This state represents the scenario where the individual has food to consume.
2. **You Don't Have Food**: This state represents case where the individual has run out of food.
3. **Neighbor Suspects You**: This state represents case where neighbour suspect you stealing his food.

# Actions

Actions are defined for each state, determining the transitions and influencing the outcomes based on the choice made:

* *Eat Own Food*: Consume available food.
* *Buy Food*: Buying food.
* *Take Neighbor Food*: Attempting to take food from the neighbor.

# Transitions

Transitions between states are determined by the actions taken. Some action in a particular state has an uncertainty of outcome:

* From **You Have Food**:
  + *Eat Own Food* always result in "You Don't Have Food".
  + *Take Neighbor Food* has a 20% chance of resulting in "Neighbor Suspects You" and an 80% chance of resulting in "You Have Food".
* From **You Don’t Have Food**:
  + *Buy Food* always results in "You Have Food".
  + *Take Neighbor Food* has a 50% chance of resulting in "Neighbor Suspects You" and a 50% chance of remaining in "You Have Food".
* From **Neighbor Suspects You**:
  + *Buy Food* transitions the individual back to "You Have Food".
  + *Take Neighbor Food* has a 90% chance of resulting in "Neighbor Suspects You" state and a 10% chance of resulting in "You Have Food".

# Rewards

Rewards are associated with transitions between states, based on the actions taken:

* **You Have Food**:
  + *Eat Own Food* leads to a reward of +1.
  + *Take Neighbor Food* could result in a reward of +1if goes to “You Have Food” or a penalty of -1 if "Neighbor Suspects You"
* **You dont Have Food**:
  + *Buy Food* results in a reward of -2.
  + *Take Neighbor Food* could lead to no change (0) or a penalty of -1 if goes to "Neighbor Suspects You"
* **Neighbor Suspects You**:
  + *Buy Food* has no reward of -2.
  + *Take Neighbor Food* can lead to a severe penalty of -5 if goes to "Neighbor Suspects You" or a reward of +1 if goes to "You have Food"

# Markov Decision Process (MDP) Model

A **stochastic Markov Decision Process (MDP)** is a mathematical model used to optimize decision-making in environments where outcomes are uncertain and depend on both the actions of a decision-maker and random events.

s defined in mathematical terms by a set of states , a set of actions , transition probabilities ), and a reward function ). Here, ), denotes the probability of transitioning to state from state after taking action , and ). represents the reward received after moving from state to state via action .

The objective in an MDP is to find a policy, a function from states to actions, that maximizes the expected cumulative reward. This is typically calculated as the sum of discounted rewards over time, formalized as where is a discount factor that weighs the importance of future rewards.

The transition probabilities determine the likelihood of moving from one state to another given a specific action, and rewards provide a numerical value for each state transition. The objective in an MDP is to find a policy that maximizes the expected sum of rewards over time.

## Initialization and Configuration

The **MDP** class is initialized with five main components: states, actions, transitions, rewards, and current state.

A screenshot of a computer program

Description automatically generated

1. **States**: Represents the different conditions that can exist within the environment as a list. In this case:

A black screen with green text

Description automatically generated

1. **Actions**: Specifies possible actions that can be taken in each state:

A computer code with green text

Description automatically generated

1. **Transitions**: Defines the probability of moving from one state to another given a specific action:

A screenshot of a computer program

Description automatically generated

1. **Rewards**: Maps each state-action-next state triplet to a numerical reward, quantifying the immediate value of a transition:

A computer screen shot of text

Description automatically generated

## Methods

* **reset()**: Resets the MDP to a random initial state.

A black background with white text

Description automatically generated

* **step(action)**: Takes an action, determines the next state from current state based on the transition probabilities, updates the current state, and returns the new state, reward, and whether the state is terminal.

A computer screen shot of a program

Description automatically generated

* **get\_available\_actions()**: Returns a list of valid actions that can be taken from the current state.

A black background with white text

Description automatically generated

## Stochastic Process Example

A computer screen shot of a program

Description automatically generated

# Conclusion

This MDP for Student Housing provides a clear model for predicting the best strategy in managing food resources and interactions with neighbors. The model defines possible states and actions, making it possible to evaluate the outcomes of various decisions under uncertainty and to strategize effectively for optimal results in a shared living environment.