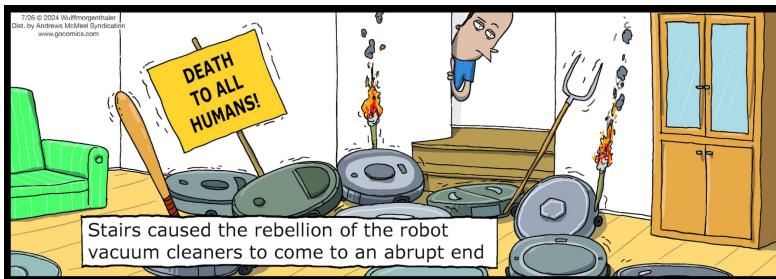


## CSCI 3202

### Lecture 27

October 29, 2025



Wumoo. <https://wumo.com>

### Announcements

- Homework 7 released today
  - Due Friday, November 7 by 11:59 pm
- Intermediate report for project due next Wednesday

### Project

- [CSCI 3202 Course Project 2025.pdf](#)
- Intermediate report due next Wednesday
- See TAs, CAs or me with questions

### Markov Models

- Covered Value Iteration last time
  - Using the Bellman equations, iterate over all states to maximize utility until they converge to the optimal values

$$U_{i+1}(s) = R(s) + \gamma \max_{a \in A(s)} \sum_{s'} P(s'|s, a) U_i(s')$$

- Once you have the optimal values, calculate the optimal policy with

$$\pi_s^* = \arg \max_{\pi} U^\pi(s)$$

- Questions
  - What is  $P(s'|s, a)$ ?

- Why do we often set  $\gamma < 1$ ?
- What is  $R(s)$ ?
- What is  $\pi$ ?
- Does  $\pi$  matter in the Bellman equations?
- What is the difference between Value and Policy Iteration?

### **Policy Iteration**

- A good comparison of Value Iteration and Policy Iteration is in Geeks for Geeks  
<https://www.geeksforgeeks.org/what-is-the-difference-between-value-iteration-and-policy-iteration/>

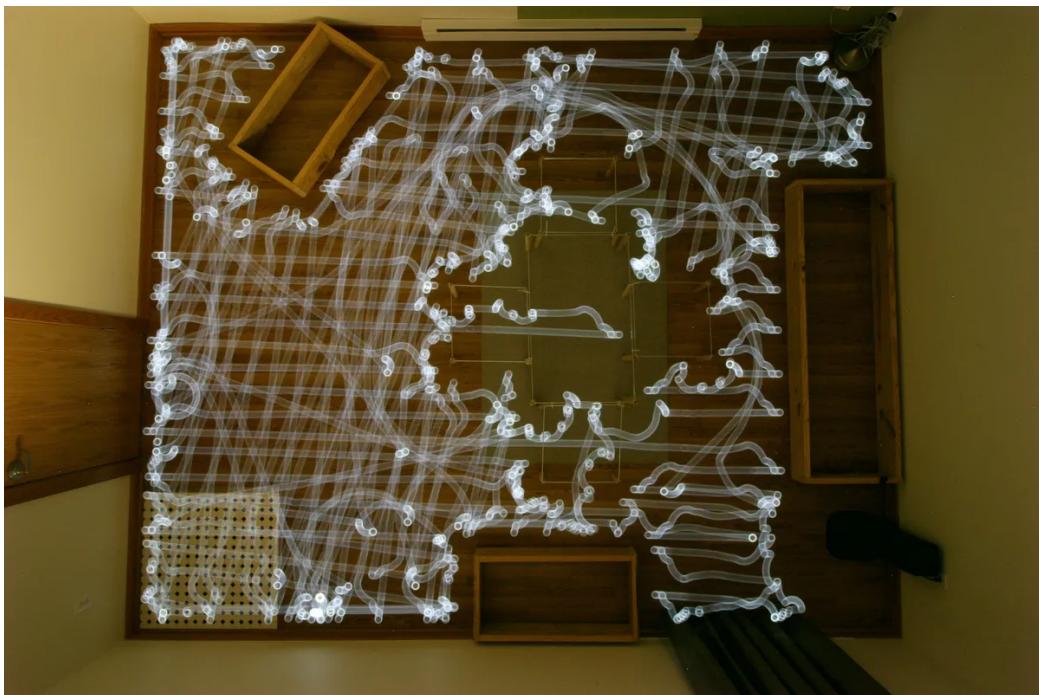
### **Reinforcement Learning and Q-Learning**

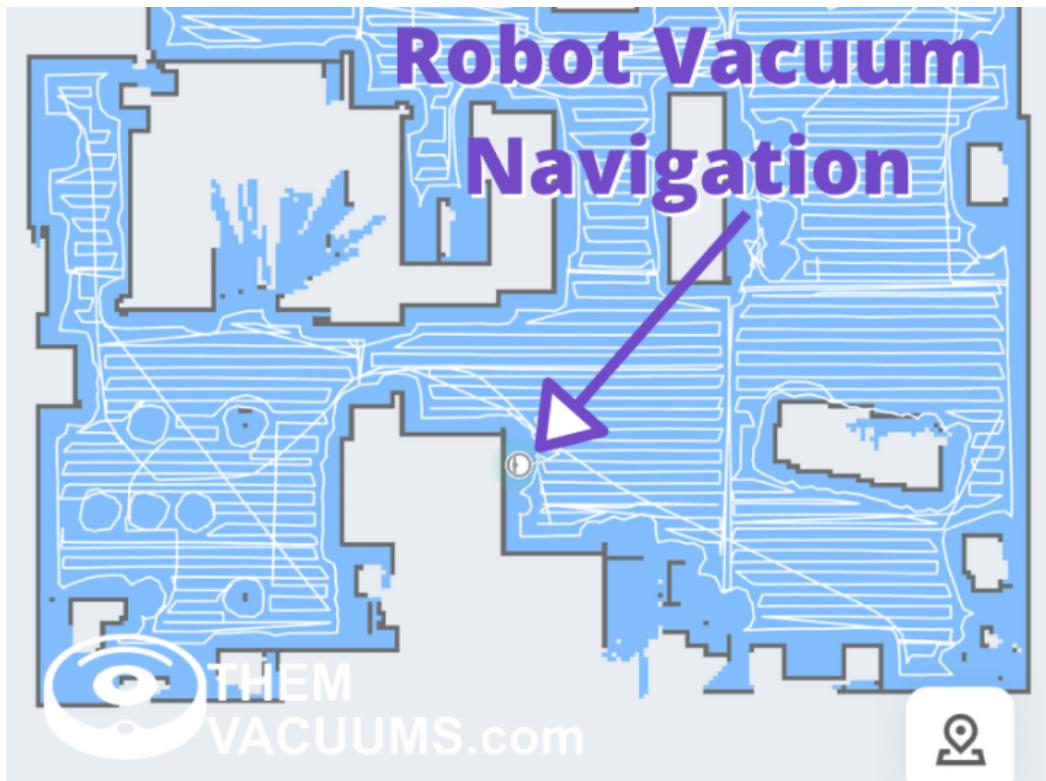
- Roomba and other Robot Vacuums
  - Typically vacuum a room at night when there is little traffic
  - Want to vacuum the whole room, not just the edges
  - Avoid obstacles and stairs
  - May map the room as they go
  - Vacuum until they run out of battery



- Assume that the Robot Vacuum uses a Markov Decision Process

- What does policy mean?
- How does it know when to stop?
- What does optimality mean with a robot vacuum?
- What could the reward function look like?
- What happens when the contents of a room move?
- What happens when someone plops something (like a coat or shoes) down in the middle of a room?





- Active vs. Passive

- Exploration vs. Exploitation
- Q-Learning doesn't require a model, but comes at a cost in terms of how many iterations it takes to create a stable utility function
- Policy Iteration Q Learning.pdf

### **Upcoming**

- Policy Iteration
- Q Learning
- Hidden Markov Models
- Quiz on Friday
  - Value iteration
  - Policy iteration